Technical Translation

Usability Strategies for Translating Technical Documentation

JODY BYRNE

Traduction Technique

Τεχνική μετάφραση

Technische Vertaling

技术转换

Technical Translation

Traducción Técnica

Traduzione Tecnica

Технически Перевод

技術的な変換

Tradução Técnica

기술적인 번역

Springer

Technische Übersetzung

TECHNICAL TRANSLATION

Technical Translation

Usability Strategies for Translating Technical Documentation

JODY BYRNE

University of Sheffield, UK

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN-10 1-4020-4652-9 (HB)
ISBN-13 978-1-4020-4652-0 (HB)
ISBN-10 1-4020-4653-7 (e-books)
ISBN-13 978-1-4020-4653-7 (e-books)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

www.springer.com

Printed on acid-free paper

Printed in the Netherlands.

Contents

Preface

Based largely on my doctoral dissertation "Textual Cognetics and the Role of Iconic Linkage in Software User Guides", this book is intended to serve as an introduction to technical translation and usability for translators and translation researchers. In this book we will look at how it is possible to improve the quality of technical translations by drawing on cognitive psychology, usability engineering and technical communication to develop skills which can be implemented during the text production stage of the translation process to ensure more usable texts. This book draws on a broad range of research and makes it accessible and applicable to an audience with a background primarily in translation although those with backgrounds in technical writing will also find the discussions of usability, cognitive psychology and usability testing useful.

Technical translation has traditionally been regarded as the poor cousin of "real" translation. Often regarded as a vocational, practical and at times rather basic type of translation, it has been largely neglected in the literature on translation theory. The work that has been done in this area has largely been restricted to terminological issues or technical issues (e.g. tools such as translation memories and machine translation, etc.) or does not fully reflect the reality of modern translation and needs to be updated (e.g. Pinchuk 1977, Sykes 1971). However, technical translation is a much more promising avenue of theoretical investigation than many suspect. Indeed, its inevitable roots in commercial translation have served to make this an even more rich and complex area than previously believed.

In recent years, the range of technical texts facing translators has grown significantly. No longer is it enough for a translator to merely understand the source text and be able to render it into a comprehensible target text. Nowadays, clients spend vast amounts of money on professional technical writers, document design and testing procedures while technical writers spend years studying how best to present technical information to users. Technical translators, on the other hand, do not necessarily receive the training and exposure to the text types which have become so prevalent or to the processes needed to create them. There is, therefore, an urgent need

to incorporate knowledge of technical communication and usability testing into the theory and practice of technical translation.

Aims and Structure of this Book

The book aims to show how to improve the usability (and consequently, the quality) of technical translations within the context of target-orientated models of translation, while at the same time recognising that there is, theoretically at least, a division of labour between author and translator. Beginning with a discussion of accepted translation theory, Chapter 1 explains that existing theories of translation do not fully consider the reality of translation as form of technical communication. Subsequent discussions make the case for a communicative approach to translation where the emphasis is very much on the target audience, rather than the original source language audience or the author.

We then examine the field of technical communication as a form of professional communication and as a "supplier" for the translation industry. Chapter 2 will explore the motivations for producing technical documents and look at the various types of documents and why they are produced. After introducing a number of fairly typical types of technical texts we turn our attention to the genre of software user guides. This genre is chosen for a number of reasons. Firstly, the ubiquity of software user guides in itself merits attention. Secondly, software guides are good examples of instructional texts, a genre which also includes training and educational materials, operating guides as well as multimedia materials. Indeed, many other types of technical text perform some form of instructional function and so, an examination of user guides will provide knowledge which can applied in a variety of contexts. The book will then set about discussing how to improve translated instructional texts using an understanding of human cognitive psychology coupled with various strategies and methods garnered from technical communication.

Chapter 3 provides a detailed examination of the fundamentals of cognitive psychology and explains how humans read, understand and learn, both in general and from texts. By understanding what it is that makes humans use and understand texts, we are better placed to discover ways of making sure the interaction between reader and text is more effective. This chapter aims to provide this grounding in human cognitive abilities and limitations before highlighting the way they affect how we read texts, particularly software user guides. The aim here is to help translators understand the

problems posed by technical instructional texts for average readers and to explore the potential for drastically improving the quality of translated technical texts using both linguistic and non-linguistic methods. Ultimately, such endeavours will ensure that texts complement the cognitive abilities of readers while at the same time, compensating for the limitations of the human cognitive system.

Having discussed human cognition and outlined the role this can play in producing better translations, the book applies this knowledge to the text production stage of the translation process. The concept of usability will be introduced as a truly indicative measure of the effectiveness of a text. By ensuring that readers can use the information in a text effectively, we make sure that the text succeeds in its primary function to instruct, or rather to educate. A detailed discussion of usability will be presented which will include definitions of usability and human factors, factors which affect usability, development processes to promote usability and key characteristics of usable texts. Usability strategies will be examined under the categories of principles, guidelines and rules.

With this theoretical basis, the next stage is to apply it in practice. A case study will be presented in which one example of a guideline along with several of its associated rules are tested as part of an empirical study conducted in order to test whether it is possible to improve the usability of translated texts using linguistic methods alone. The case study simulates the text production stage of the translation process and examines whether Iconic Linkage, the process of replacing non-isomorphic but semantically identical segments of text with isomorphic formulations can improve usability. Iconic Linkage will be defined and discussed. A range of examples are provided to illustrate the concept in a multilingual context. A detailed description of the rationale, preparations, methods and results of the empirical study will be presented and discussed.

The book will conclude by evaluating the previous sections and examining other ways in which textual cognetics can be used in translation theory and practice.

Using this Book

This book is aimed at a number of potential audiences, each of which may have different needs and expectations when it comes to reading this book. The following list is intended to help you find information on individual subject areas:

- Translation theory and technical translation: Chapter 1
- Cognitive psychology: Chapter 3

- Experimental methods: Chapter 5
- Usability: Chapters 3, 4 and 5
- User guides: Chapter 2
- Reading: Chapter 3
- Readability: Chapter 2
- Technical communication: Chapter 2
- Writing strategies: Chapters 2 and 4

Chapter 1 provides a general overview of translation in general and technical translation in particular. It does not represent a comprehensive review of all the literature on translation but it does provide a good overview of the main themes in translation as they relate to technical translation.

Chapter 2 examines technical communication as a creative and technical discipline. It is useful for understanding why technical documents are produced, how they are produced and by whom. This chapter also describes user guides in detail and describes the various requirements which need to be met for a user guide to be regarded as successful and effective.

Chapter 3 is of benefit to those wishing to understand the cognitive processes which allow humans to perceive information, read documents and learn new information. It provides an introduction to human cognition and explains how this mechanism facilitates learning and reading.

Chapter 4 deals with how we can use an understanding of human cognition, translation and technical communication to engineer user guides so that they are as usable and effective as possible. This chapter looks at usability in detail and discusses various criteria which can be used to quantify

usability. After introducing the notion of usability principles and rules, the chapter presents Iconic Linkage as one possible method for improving usability in user guides. Read in conjunction with Chapter 3, this chapter provides a solid basis for understanding usability in general and with regard to user guides.

Chapter 5 draws on all of the preceding chapters and sets out to implement measures to improve usability and then assess the effectiveness of the measures. This chapter provides a detailed introduction to usability testing and evaluates a wide range of literature on the subject. The chapter culminates in a practical usability study which draws on a range of theoretical and practical sources. The results of the study are assessed, evaluated and discussed.

Chapter 6 attempts to draw conclusions from the preceding chapters and discuss the impact of usability on technical translation as well as implications for future research.

Acknowledgements

In preparing this book several people have provided invaluable help and support for which I am truly grateful. First of all I would like to thank Dr Dorothy Kenny, my doctoral supervisor at Dublin City University (DCU), whose patience, attention to detail and ability to make sense of my often convoluted thought processes played an enormous role in the completion of my research.

My gratitude also goes to the following people:

- Danielle Barbereau and my colleagues in the *Modern Languages Teaching Centre* at the University of Sheffield for their patience and support while I completed this book.
- My friends and former colleagues at DCU for their help and support.
- The School of Applied Language & Intercultural Studies at DCU for the financial assistance which helped support my doctoral research.
- Lena Hamaidia, Dorothy Kenny and Mona Baker who very generously agreed to read various chapters and offered valuable advice and suggestions. Ultimately, however, any shortcomings in this work are the result of my own stubborn nature.
- Nick Brooker for his superb work on various illustrations in this book.
- My mother Ellen for being an outstanding mother, friend and inspiration.
- My sister Sinéad for being an amazing sister and friend, and for being a model of courage in the face of adversity.
- My grandmother, Éilis, for her love, support and timely cups of tea while I wrote my dissertation.
- Willie Bryans for his support and invaluable assistance in conducting my empirical research.

Finally, I am eternally grateful to Janice for her love, patience, understanding and support throughout my research and her ability to make me see the funny side of just about everything. Thank you!

Chapter 1
Technical Translation

Technical translation has long been regarded as the ugly duckling of translation, especially in academic circles. Not particularly exciting or attractive and definitely lacking in the glamour and cachet of other types translation, technical translation is often relegated to the bottom division of translation activity and regarded as little more than an exercise in specialised terminology and subject knowledge. Indeed, these factors, particularly subject knowledge, have in some quarters led to technical translation being feared and loathed, like a modern-day barbarian of the linguistic world.

That technical translation has traditionally been regarded as the poor cousin of "real" translation in the literature is clear. This vocational and industrial type of translation has been largely neglected in the literature on translation theory. This is supported by an enlightening survey by Franco Aixelá (2004) who reports that out of 20,495 publications listed in the BITRA[1] multilingual bibliography of translation research only 1,905 or 9.3% addressed technical translation. Literary translation, on the other hand, is the subject of some 4,314 entries accounting for 21% of the total number of entries despite its niche status in professional practice.

The work that has been done in this area has largely been restricted to terminological issues or technical issues (e.g. translation memories or machine translation, etc.) or needs to be updated to reflect the modem realities of technical translation (e.g. Pinchuck 1977, Sykes 1971). However, technical translation is a much more promising an avenue of theoretical investigation than many suspect. Indeed, its inevitable roots in commercial translation and technical communication have served to make this an even more rich and complex area than previously believed.

---

[1] BITRA is the *Bibliography of Interpreting and Translation* which was created by Javier Franco in the Department of Translation & Interpreting at the University of Alicante. This useful web-based resource can be found at http://cv1.cpd.ua.es/tra_int/usu/buscar.asp?idioma=en

The aim of this chapter is to challenge some misconceptions about technical translation and describe the reality of this form of translation. I will also try to relate technical translation to some of the more common theories of translation. There are two main reasons for this. Firstly, to show that technical translation is worthy of theoretical study and secondly to show that technical translation, like other specialised types of translation, does not fit neatly into any one theory or approach and that there is, as yet, no adequate explanation of technical translation as an activity.

The Importance of Technical Translation

It has been estimated that technical translation accounts for some 90% of the world's total translation output each year (Kingscott 2002:247). This is unsurprising given the importance attached to the availability of technical information in a variety of languages, motivated partly by the increasingly international focus of many companies and partly as a result of legislation such as Council of the European Union Resolution C411 (1998a), EU Directive 98/37/EC (Council of the European Union 1998b) and Council Directive 93/42/EEC (1993) and international standards such as EN 292-2: 1991 and EN 62079: 2001 to name just a few. These represent just some of the various laws, directives and regulations across the world that require the provision of comprehensive, accurate and effective technical documentation in a variety of languages. Coupled with increasing international cooperation in scientific, technological and industrial activity, it is clear to see why technical translation is one of the most significant employers of translators.

Yet despite the overwhelming demand for and importance of technical translation, there are several stubbornly persistent myths about technical translation's importance, nature and role both in industry and within academia.

Some Misconceptions

Before we examine technical translation in greater detail and try to relate it to various theories of translation, it would be useful to look at what we mean by "technical translation" and contrast some misconceptions about technical translation with the realities of what it means to be a technical translator.

Technical translation includes economics, law, business etc. In reality, "technical" means precisely that, something to do with technology and technological texts. Just because there is a specialised terminology, doesn't make something technical. In discussing technical translation it is useful to make the distinction between specialised and technical translation. For example, religion has a very specific terminology and very definite conventions, styles and document structures but it is never regarded as "technical". The tendency among certain theorists to include LSP texts such as legal, financial and economic texts within the field of technical translation is less than helpful not least because each area has its own unique characteristics, requirements and constraints. Simply because a field or subject area has unique or specialised terminology does not make it technical. This is not to say that financial translation, or indeed legal translation, do not deserve to be studied in detail as areas in their own right, in fact there are a number of extremely useful books on these areas such as Alcaraz & Hughes (2002), but rather that they will not be discussed here. Instead, this book will take as its basis a definition of technical translation that has its roots in the translation industry and indeed industry as a whole, namely, that technical translation deals with technological texts. Or more specifically, technical translation deals with texts on subjects based on applied knowledge from the natural sciences.

Technical translation is all about terminology. This particular misconception is not unique to those uninvolved in technical translation. A surprising number of people within technical translation share this belief Pinchuck (1977:19), for example, claims that vocabulary is the most significant linguistic feature of technical texts. This is true insofar as terminology is, perhaps, the most immediately noticeable aspect of a technical text and indeed it gives the text the "fuel" it needs to convey the information. Nevertheless, Newmark (1988) has claimed that terminology accounts for at most just 5-10% of the total content of technical texts yet there is a disproportionate amount of attention devoted to terminology and lexical issues in technical translation. A simple subject search for "technical translation" on the BITRA bibliographic database reveals that more than half of the 150 entries found relate to terminological or lexical issues.

What makes this even more surprising is the fact that in many fields of science and technology, the terminology is remarkably similar to the extent that separate, specialised dictionaries are frequently unnecessary. Indeed, Fishbach (1993 and 1998) points to the quasi-conformity of medical terminology thanks to the common origins in Latin and Greek. So, depending on the particular language pairs, a translator should have less trouble locating appropriate specialised terms in the target language than with

non-specialised, general terms. Similarly, in computing and IT, the terminology is largely uniform thanks, in part, to a predominance of English in the creation of new terms and partly to the proliferation of proprietary terms and the availability of terms from software companies, e.g. the Microsoft glossaries which are available in every language into which Microsoft's products have been localized.

However, perhaps even more important than terminology is actually knowing how to write the texts. Translators need to produce texts which are identical to those produced by technical writers working in the target language (Fishbach 1998:2). Failing to comply with target language text conventions can undermine the credibility of the text, the author and the information in the text. O'Neill (1998:72) claims that "there is no substitute for a thorough knowledge of the target language". In order to do this, it is necessary to look to technical writing and this is not something many translators have the opportunity to do, either as part of their training or as part of their own efforts to improve their skills.

According to Lee-Jahnke (1998:83-84), there are three things that are essential in order to learn how to deal with scientific and technical texts:

- know the text structure in the different languages
- know the LSP for the area
- know the subject area

Style doesn't matter in technical translation. This is, perhaps, one of the more irritating misconceptions for technical translators because it is so completely unfounded and implies that technical translators do not have the same linguistic and writing skills as other types of translator. Perhaps the problem stems from differing opinions of the nature of style and the popular belief that it relates exclusively to literature. If we look at style from a literary point of view, then it does not have any place in technical translation. But if we regard style as the way we write things, the words we choose and the *way* we construct sentences, then style is equally, if not more, important in technical translation than in other areas because it is there for a reason, not simply for artistic or entertainment reasons. As Korning Zethsen (1999:72) asserts, literary texts "do not hold a monopoly on expressivity and creativity". To illustrate this, consider a leaflet containing instructions for using a product. The limited space available requires both the author and translator alike to express information in a way which is sufficiently clear, simple and concise so as to allow readers to understand the information completely and quickly but which nevertheless conveys all of

the necessary facts. In comparison, consider a poem where an author may purposely choose stylistic devices, structures and metaphors which will make the meaning ambiguous and leave it open to several interpretations so as to add to the readers' enjoyment of the poem. Both situations will require the use of stylistic and expressive language in order to achieve the desired effects although these approaches may be at opposite ends of the stylistic spectrum.

In many cases, the importance or even existence of style in technical texts goes completely unacknowledged, due largely to the belief that because technical language is functional, it must be "plain" and stripped of any form of style or linguistic identity. In reality, however, technical translation is a highly complex endeavour and style is one of its most important facets. For this reason, this book will take as its basis the concept of style and its application in technical translation. This book will show that style, which has been regarded at best as a way of ensuring compliance with target language norms, can actually have much more profound effects on the quality of technical translations.

Technical translation is not creative; it is simply a reproductive transfer process. While technical translation "is undoubtedly more restricted in range than aesthetic translation" it is much too easy to overestimate and exaggerate its apparent simplicity (Pinchuck 1977:20). But in order to convey information in an appropriate and effective way, technical translators have to find novel and creative linguistic solutions to ensure successful communication. That this task is often hampered by a restricted vocabulary and stylistic constraints merely makes the achievement all the more impressive.

You need to be an expert in a highly specialised field. There is a common belief that in order to be a good technical translator, you need to be an expert in a highly specialised field and you can't specialise in more than one or two subject areas. But the reality is that, armed with a good and solid understanding of the basic principles and technologies, many technical translators can, in the words of Robinson (2003:128) "fake it". He says that "translators… make a living pretending to be (or at least to speak or write as if they were) licensed practitioners of professions that they have typically never practiced." They are like actors "getting into character".

However, lest technical translators be branded a bunch of scurrilous charlatans who deceive unwitting clients we need to put Robinson's comments into perspective. The notion of *pretending* to be an expert means that the translator should have enough subject knowledge either to know how to deal with the text or to be able to acquire whatever additional information is needed. Researching a new subject area for a translation is

always easier when you know *at least something* about it compared to when you know nothing at all. It is, therefore, essential that translators have excellent research skills, make full use of parallel texts and have a very good understanding of general scientific and technological principles. Technical translators need to "impersonate" the original author who is generally, though not always, an expert in a particular field and they need to write with the same authority as an expert in the target language. So in this case, the real challenges for the technical translator are to be able to research subjects and to have expert knowledge of the way experts in a particular field write texts. We can summarise the essential areas of expertise for technical translators:

- subject knowledge
- writing skills
- research skills
- knowledge of genres and text types
- pedagogical skills

With regard to subject knowledge, O'Neill (1998:69) says that medicine, nursing, dentistry etc. all share some subject knowledge, e.g. physics, biology, pharmacology etc. Consequently, what are normally regarded as quite separate specialisms frequently share the same basic knowledge. The same can be said of various other types of technological area: engineering, mechanics, construction etc. They all share more or less the same core knowledge.

When it comes to specialised subject knowledge, it is a truism that you cannot master every area so it's probably best to get a good foundation in the more generic, transferable subject areas and rely on this together with an ability to research new areas to deal with unfamiliar subjects. So for a technical translator, gaining a good understanding of the basics of science and technology can provide a good basis for diverse applications within technical translation.

Technical translation is all about conveying specialised information. This is not entirely true, of course the main concern for technical translators is to make sure that information is conveyed accurately but they are also responsible for ensuring that the information is presented in the correct form, that it is complete and that the information can be used correctly and effectively. The translator's responsibilities encompass many of those of the technical author and that may involve quite radical changes. Technical

translation involves detailed knowledge of the source and target cultures, target language conventions, text type and genre conventions, register, style, detailed understanding of the audiences; whether translators realise it or not, an understanding of how people learn and use information. Subsequent chapters in this book, particularly Chapter 2, will show that merely presenting information to readers is not enough. Readers must be able to assimilate the information with as little effort as possible. The information presented in technical documents is a means rather than an end and it should not add to the work of readers.

The Reality of Technical Translation

Having looked at some of the more prevalent and serious misconceptions of technical translation, it's time to look at the truth behind technical translation. Our aim here is not to provide an exhaustive and comprehensive treatment of the area but to give a good general overview and insight into the area. This section seeks to describe the reality of technical translation in a professional context, to describe what it is that technical translators do and what factors affect their work.

Scientific vs. Technical Translation

One of the greatest fallacies when discussing technical translation is to somehow lump it together with scientific translation, or worse still, to use the two terms interchangeably. Throughout the literature on translation, in the frequently brief references to technical translation we see the expression *scientific and technical translation,* where, out of convenience perhaps, authors see no problem in treating these two siblings as conjoined twins or even as the same person. This fundamental error serves only to confuse the issue because scientific and technical translation are not the same and as such, cannot be compared equally.

Despite the obvious connection between the two, i.e. they both deal with information based, to varying degrees, on the work of scientists, scientific translation is quite distinct from technical translation. Certainly, they both contain specialised terminology and, on the surface, deal with complicated scientific subject matter (to an extent) but it is all too easy to overestimate these apparent similarities at the expense of other, more compelling, differences.

One of the easiest ways of disambiguating the matter is to look at the words themselves: *scientific* and *technical. Scientific* relates to science

which is defined by the Chambers Dictionary as "knowledge ascertained by observation and experiment, critically tested, systematised and brought under general principles" (Chambers 1992). *Technical* relates to technology which is defined as by the Concise Oxford English Dictionary as "the application of scientific knowledge for practical purposes". Thus we can say that scientific translation relates to pure science in all of its theoretical, esoteric and cerebral glory while technical translation relates to how scientific knowledge is actually put to practical use, dirty fingernails and all. The differentiation between scientific and technical translation is also acknowledged by the information sciences. Pinchuck (1977:13) points out that even in libraries, pure sciences are classified under 5 while applied sciences, i.e. technological material, are shelved in their own place under 6.

Scientific and technical translation, therefore, is a generic term which is used to refer to pure science, applied scientific research and technology. But it is not just the subject matter that distinguishes scientific from technical translation. Technical translation (and technical communication, which will be covered later on) can be characterised at a basic level on the basis of:

1. subject matter
2. type of language
3. purpose

So we can, for example, translate a scientific paper which deals with the concept of electromotive force and the effects of currents passed through conductors, complete with formulae, hypotheses, discussions and calculations or we can translate an installation guide for an electrical motor. Both texts are based on the fact that if you pass an electrical current through a piece of wire, a magnetic field is created which exerts a force acting at right-angles to the wire. The difference is the way in which the knowledge is used and presented. And this is a fundamental difference between scientific and technical translation and one which also affects the type of language used in the texts.

In our scientific paper on electromotive force, the goal is to discuss, explain, justify, impress, convey, convert and possibly entertain. An author will use the full extent of the language, within specific conventions and norms, to present the information in an interesting, serious and impressive way. In some cases, these texts even border on the literary, using the same devices and strategies as a short-story writer or novelist. Scientific language

can be quite formal[2] (think of journal papers) and will often have considerable range, just like a literary text. Such texts will also see the use of various rhetorical strategies, Greek and Latin terms and expressions as well as various affixes and compound terms.

The following examples illustrate the type of literary language use which can be found in scientific texts and even within a single text. In the introduction to a detailed book on astrophysics, Schatzman and Praderie (1993:1) paint a rather picturesque scene:

> In the splendour of a moonless night, far from the pollution of the sky by artificial lighting, the first revelation is that of the stars.

The next sentence is taken from a book discussing the scattering of X-rays in gases, liquids and solids where the author uses rhetorical questions, among other strategies, to enrich the text.

> How is the sudden decrease of μ in Fig. V.5 explained by resonance? (Hukins 1981:47)

In discussing the origins of the universe and the *Big Bang* theory, Goldsmith (1995:68) uses the expression "*tough little devils*" as a humorous way of explaining the nature of helium nuclei:

> This tiny fraction results from the characteristics of helium nuclei, tough little devils that cannot easily be made to fuse into larger nuclei, because no stable nuclei exist with either five or eight nucleons (protons or neutrons).

An installation guide, on the other hand, is written to help someone do something. The aim here is to convey the information an engineer needs in order to install, connect and commission the motor. Consequently, the language used will reflect this: simple, unambiguous, concise, and, for want of a better word, unremarkable. The aim here is not to entertain the reader.

[2] Popular science books and magazines which form a sub-class of scientific texts tend to have a less formal and possible more journalistic tone but they are still capable of switching between a jovial, friendly style and a more formal "scientific" tone.

People generally do not settle down with a glass of wine to spend the evening reading installation guides for fun. Pinchuck refers to technical language as "workshop language"[3], which is somewhere between scientific and general language. It is less regulated, less literary and even colloquial on occasion but always strictly functional.

Scientific texts will be conceptually more difficult and will be more abstract than other types of text. They will, however, have more standardised terms which are easier to look up and they are likely to be better written than texts on other levels. Technology-based texts will be more concrete, will contain less information in more space, they will be more colloquial and will feature concepts which are easier to understand. In addition to this, there will be products and processes in the external world which can be referred to. In other words, technical texts can rely on world or background knowledge to a greater extent. (Pinchuck 1977:218-219).

The Aim of Technical Translation

This leads us on to examine the aim of technical translation. While the preceding discussion would lead us, quite justifiably, to say that the aim of technical translation is to transmit technical information, this would be just half of the story. Although it is true that technical texts are utilitarian (Pinchuck 1977:18) and are intended to serve a relatively finite purpose, namely to clearly present information to the target language readers, there is more to technical translation than simply transmitting information. Instead, the challenge for technical communicators is to ensure that all of the relevant information is indeed conveyed but also that it is conveyed in such a way that the readers can use the information easily, properly and effectively. Indeed, this aim is precisely the same as that of technical writing, which, rather unsurprisingly, forms the basis for technical translation in that it supplies the raw materials for translation activities.

[3] Pinchuck (1997:163-164) is eager to stress that *workshop* is not meant in any derogative sense. Rather, it reflects one of the traditional areas where scientists work and as such is equally as valid as "scientific". In any case, he offers an interesting discussion of the differences between scientific and technical language.

A Communicative Service

In previous paragraphs we referred to technical translation as a communicative service. Indeed, this is reflected in the following quote from Sykes (1971:1):

> Practical translating... is a service industry. The value of the service provided by... the translator, depends here, primarily, not on the effort which went into its preparation, its literary merit, its quality of presentation, production and reproduction, etc., but on its gap-bridging capacity, its message and content, its scientific or commercial utility to the requester.

The purpose of technical translation is, therefore, to present new technical information to a new audience, not to reproduce the source text, *per se,* or reflect its style or language. Technical translation is a communicative service provided in response to a very definite demand for technical information which is easily accessible (in terms of comprehensibility, clarity and speed of delivery).

As a communicative service or act, translation inevitably involves people. At the most basic level, we can say that this communicative act involves the author, the translator and the reader. This is the standard way of examining the translation process. However, this is far too simplistic for anything other than the most basic discussions.

In reality, several parties are involved in the process. This has been addressed by authors such as Vermeer and Nord but it is necessary here to emphasise the important roles played by the various stakeholders. Conscious of the fact that the majority of non-literary documentation is not produced as a result of some desire or intention on the part of the author, Nord (1991:42ff) distinguishes between the sender and text producer (author). Using software user guides as an example, we can say that the sender would be the software company who actually develops and sells the software. As part of the product development process, the company instructs one of its employees, a technical writer, to produce a comprehensive user guide to help users learn to use the software. Thus, the text producer in this case is the technical writer working for the software company.

An interesting point is made by Nord (1997:21) who claims that unless a source text was produced specifically to be translated, the source text author has no immediate involvement in the translation process. This poses something of a problem because in the current climate with multilingual, multinational companies and legislation requiring documents in multiple

languages, it is difficult to imagine a case where it does not occur to someone during the production of documentation that it may be translated at some stage, even if it is just a distant prospect. Still, it is hard to see why the author would be involved in this process unless the translator needed to clarify an ambiguous part of the text. Again, with the growing dependency on freelance translators, it is unlikely that a translator would have any direct or indirect communication channel with the author. A peculiar situation indeed and, perhaps, worthy of further investigation in the future.

We can see from this, however, that the stakeholders in the technical translation process are not simply the author, translator and reader. If we examine the practical circumstances surrounding the production of multilingual documentation, we can propose a number of other parties:

- Document Initiator
- Writer/Text Producer
- Translation Initiator
- Translator
- User

The *Document Initiator* is the person or entity responsible for the production of the original source language document. In the case of product documentation, this is invariably the company that manufactures the product. The Document Initiator's aim here is to have documentation that will help users learn how to use the product effectively, safely and efficiently. This can be motivated by a genuine desire to help users, to provide a "complete" product, to improve the company's competitiveness and reputation, to reduce calls to technical support or simply to comply with legal requirements.

The Document Initiator then instructs the *Writer* (who may be in-house or a freelance contractor) to produce the documentation. The Writer either receives a detailed brief from the Document Initiator or develops a document specification on the basis of interviews with the Document Initiator and/or the Writer's own expertise. Working to this specification, the Writer gathers information from a variety of sources such as white papers, project outlines, hand-written notes, product specifications and marketing materials and processes them to produce a text which then goes through a number of iterations which is tested at each stage for content, clarity and ideally usability. The completed document is returned to the Document Initiator and distributed to customers. By referring to the *sender* and the *text producer* (Nord 1991:42ff) also acknowledges the need to differentiate

between the person who actually writes a text and the person or entity who orders its creation.

The *Translation Initiator* is the person or entity responsible for starting the translation process. This is generally the Document Initiator but it can be argued that the Translation Initiator may be a different department or manager within the same company. The motivations for the translation process are similar to those for the document production process, i.e. a desire to enter into new markets where the documents serve not only as a way of gaining customers but also as an "ambassador" for the company and its products. Translations may be motivated by a need to comply with legal requirements such as the Council of the European Union *Resolution C411* which states that "customers are entitled to manuals produced in their own language" irrespective of where the product was made (Council of the European Union 1998:3). The Translation Initiator sends the document and some form of instructions to the Translator.

While strictly speaking the *Translator* is the next step, in reality this is not always the case. Frequently, texts will be sent to *translation agencies* or *localization vendors* who will then send the document to a translator. It goes without saying that this additional link in the chain can present both problems and benefits for the Translation Initiator and the Translator. Problems arise from the fact that this is another stage in the communication process and as such, any instructions or requirements specified by the Translation Initiator may well be "watered down", misinterpreted or not passed on by the agency or vendor. Of course, this depends on the systems and processes the agency has in place as well as on the personalities and communication skills of the project managers and administrators involved. It is not unheard of for a project manager to simply send a text out for translation by a freelancer without passing on any form of instruction, assuming instead that the translator will know what to do with it. On the other hand, certain agencies have robust processes in place to ensure that translators are given detailed information such as whether the text is for publication or for information purposes, whether specialised terminology or style guides need to be used and so on. However, for the sake of clarity and simplicity, we will assume that the instructions from the Translation Initiator have been conveyed perfectly to the Translator and that there is no cause to suspect that ambiguities have been introduced or instructions lost.

Translators can be either staff (working for the Document/Translation Initiator), in-house (working in-house for an agency or vendor) or freelance and they are responsible for producing a foreign language version of the original document. Unfortunately, practices regarding translation briefs or instructions for producing the translation are far from consistent and

frequently problematic. In addition to the factors discussed above, the problem is that many Translation Initiators simply do not have the necessary knowledge or experience to realise that they should give some form of brief when commissioning translations. Some simply regard this as the translator's responsibility. For instance, some might argue that you would not give a mechanic a detailed checklist and instructions for servicing your car. Instead they would expect that, having described the problem in general, the mechanic would be able to establish what the problem was and would know how to fix it. The same thinking applies when some clients think of translators. At best, many clients will simply specify that the document is for publication purposes, should use company's own terminology and should "read well". In this regard, in-house and staff translators fare a little better than their freelance counterparts.

Even in the absence of a translation brief, an in-house translator will have access to various experts, sometimes even the original author as well as access to a variety of existing documentation in the form of previous versions of the document and parallel texts. Freelancers, unfortunately, do not always have this luxury although the more conscientious will pester the client or agency for this material. More often than not, translators have to rely on previous experience and knowledge of the document conventions for both languages. So, using a knowledge of the two languages, including the methods for producing effective texts in the target language, subject knowledge and their perception of the audience's needs, expectations and requirements while at the same time ensuring that the client's wishes are met along with the legal requirements for documentation, the translator produces the foreign language version of the text. The translation is then made available to the User, who represents the final stage in the process.

It is the *User* who, according to Holz-Mänttäri (1984:111), is the decisive factor in the production of a target text. Where technical documents are translated, there are two sets of users: the source language user and the target language user. Although both users are on the surface quite different in terms of their needs, i.e. documents in the respective languages, they share certain macro-aims, e.g. learn how to use the product. However, this may need to be achieved in different ways. This is where the technical writer and the technical translator share a common purpose to determine what their respective users want and how to achieve this. In order to translate effectively, a translator needs to fully understand and know the users (Reiss & Vermeer 1984: 101).

The user is the real reason the original document was produced in the first place and subsequently translated. What the user wants from a translation has been subject to much debate, argument and hand-wringing among

translation theorists. Some argue that the target language reader may be interested in gaining an insight into the culture of the source text culture or the source language itself. Others would argue that the target reader wants to experience the text in the same way as the original audience did. In certain circumstances this may be true. However, in the case of technical translation, all readers are concerned about is getting the information they need and being able to understand and use it effectively in order to do something else, usually some task relating to their day to day work. This may sound rather absolutist and dogmatic, but in reality many people treat translations not as translations but as original target language texts. This is unless, of course, there is some quality issue within the text which identifies it as a translation in which case the user will probably lose much, if not all, trust in the text. In any case, the translation needs to function in precisely the same way as any other text in the target language. Readers are unlikely to show mercy to a translation that is obviously a translation just because it is a translation. This serves only to distract them from their primary concern: finding the information they need in the document and using it.

With this general overview of the key stakeholders involved in technical translation, we now need to take a closer look at the role of the translator. After all, it is the translator who facilitates this extension of the communication process.

The Translator's Role

Given its central position in the entire translation process, the role of the translator is, understandably, more complex than the other participants and as such merits closer investigation.

Much like the source language technical writer, the translator's primary job is to communicate information by means of a text. This aim supersedes any desire or intention to transfer the source text into the target language. As Robinson (2003:142) maintains, "translators don't translate <u>words</u> they translate <u>what people do with words"</u>. In this case, we have the added complications presented by the fact that we are dealing with two different

languages, a fairly rigidly prescribed framework within which to produce the target text (i.e. the translation brief and the source text[4]).

However, like the technical writer, the translator uses information from a variety of sources, not just the source text, to produce a target text which is effective and which performs the desired communicative function.

In this sense, the translator becomes the *intercultural* or *cross-cultural technical writer* referred to by Göpferich (1993) and Amman & Vermeer (1990:27). But this isn't just an excuse for loose and overzealous translations by translators who are dissatisfied with their role as *mere* translators and yearn for the "power" of an author. Quite the opposite in fact! The sign of a good technical translator is the ability to do **some** of the things a technical writer does to make sure that the person who ultimately reads the text can do so with relative ease and that whatever tasks the reader needs to perform, are easier having read the text.

However, in practical and professional terms, the actual work of a translator is still somewhat unclear and not fully understood. In the words of Mossop (1998:40) there are "no systematic observations, or even self- descriptions, of how professional translators proceed when they translate". This is indeed true. We may speculate as to what is actually involved in translating a text and how it is done by a translator but in terms of what a translator actually does and when, we are still guessing to a large extent. Mossop describes the process of translation as a process consisting of 5 tasks performed over 3 phases of translation production:

- Phase 1: Pre-drafting
- Phase 2: Drafting
- Phase 3 Post-drafting

[4] I am conscious of the need not to elevate the source text to such heights that it dominates and determines the translation process. The emphasis here is, and will remain, firmly on the needs of the target audience and not on the source text or author. Having said this, there is a limit, as yet intangible, to what a translator can do and how far a translator can deviate from the source text. A useful way of thinking about this is Nord's notion of "function plus loyalty" as part of her approach to Skopos theory (see page 38) although Vermeer has reservations about introducing subjective, value judgments into Skopos theory (Vermeer 1996:83).

During the course of these phases, translators perform the following tasks, although it is not clear, according to Mossop, how the tasks are distributed. It is possible that the tasks are performed sequentially or in parallel.

- Task 1: Interpret the source text
- Task 2: Compose the translation
- Task 3: Conduct the research needed for tasks 1 and 2
- Task 4: Check the draft translation for errors and correct if necessary
- Task 5: Decide the implications of the commission. In other words, how do the intended users and uses of the translation affect tasks 1 to 4?

While Mossop presents these phases and tasks as a description of practical translation as a whole, they are easily applied to technical translation, although it is possible that the distribution of tasks and stages is even more difficult to pin down. Nevertheless, this description serves to give a useful overview of the stages involved in producing a technical translation because it acknowledges the need for translators to conduct research so as to understand not just the text but also the subject while at the same time ensuring, by means of revisions and corrections, that the text conforms to target language norms and target audience expectations.

Technical Translator or Technical Communicator?

That the lines separating the role of technical translator and technical writer have become somewhat blurred is inevitable. What's more, this fact is gaining greater recognition in the wider "communication" community. Various professional societies for technical communication in Europe and in the United States - such as the Institute for Scientific and Technical Communicators and the Society for Technical Communication - specifically include translators in their definitions of technical communicators.

Not only do both camps deal in the same currency, i.e. technical information in texts, they also share several key tasks and activities. Perhaps first and foremost, technical writers are, to a certain extent, not unlike translators in that they need to "translate" what Van Laan and Julian (2002:18) call "geek-speak" into clear and understandable English. As mentioned previously, a technical writer gathers information from a variety of sources including documents that were produced by and for experts such as programmers and engineers. With this come the inevitable infelicities of style, excessively technical content or indeed missing information. This information

needs to be rebuilt, reinterpreted, remodelled and restructured so that it can be understood and used by the reader. Likewise, the translator needs to transform information from a form which was produced by and for speakers of the source language into a form which can be understood by the target audience. This is achieved by editing, rearranging, adding and even removing information.

Adding and Removing Information

Göpferich (1993:52) says that sometimes it is quite necessary to add additional information to a text to help ensure that the information and text are as usable and effective as possible. Similarly, some information needs to be omitted, condensed or made implicit because it is of less importance for the target audience, because it may not apply to the target audience's particular situation or because it may result in confusion. Put simply, sometimes information must be sacrificed in order to protect the integrity of the communication.

Pinchuck (1977:206-207;210-211) maintains that a text should give readers just enough information for their purposes; no more, no less. Too much information, like too little information can lead to confusion, stress and unnecessary effort on the part of readers. The rule of thumb is that if a user does not need to know something at a given moment, don't tell them about it. But this can be abused, either intentionally or unintentionally. In her discussion of medical translation, O'Neill (1998:76) claims that doctors-turned-translators (who have much more subject knowledge than traditional translators) tend to edit, reduce and summarise texts to make them better. However, the problem here is that unless this is specifically what the reader wants and needs, the translator is committing quite a serious error.

This type of intervention also manifests itself in the formulation and editing aspects of the translation process. In the case of poorly formulated source texts, this requires the technical translator to intervene whenever necessary in order to reword, edit or present information in the best way for the reader. According to Sykes (1971:6) the translator should "not feel compelled to perpetuate the more sinful omissions or commissions of his [sic] author". Sykes goes on to advise technical translators to "look out for unnecessary verbiage (including padding)" as it allows the translator to "re-phrase rather than paraphrase". The opposite is true of excessively concise texts (Sykes 1971:10). In fact, Ramey (1998) discusses how sentences that have been overly condensed can result in Escher effects, or sentences that can have multiple meanings that take more than a little detective work on the part of the reader to decipher.

Similarly, the translator may need to intervene in the flow of information within a text. This is particularly important in the case of user guides or instructions for use where instructions in the text must match the actual sequence of actions the reader needs to perform. This is also useful for the purposes of ensuring the appropriate register in the target text. Gerzymisch-Arbogast (1993) discusses how the sequencing of information in a text reflects its register or author-reader contract and how the register for a particular document type may need to change in translation.

Sometimes, too, changes to the sequencing of sections in a document are needed because of cultural norms relating to the structure of a particular type of document. For example, a user guide for an electrical appliance in German might typically start with an explanation of the parts, then information for disposing of the product after use followed by an introduction. In English, the user guide for a comparable (or even the same) product might start with an introduction followed by a quick start tutorial.

There is a very definite legal requirement on technical writers and in turn, on technical translators to ensure that documents contain enough information to ensure the safe operation of products and devices. European directives such as Directive 98/37/EC of the European Parliament place the onus on manufacturers (and the writers and translators working on their behalf) to predict possible hazards and to document them in instructions and product documentation. The directive states that:

> 1.2.2. Principles of safety integration
> (c) When designing and constructing machinery, and when drafting the instructions, the manufacturer must envisage not only the normal use of the machinery but also uses which could reasonably be expected. The machinery must be designed to prevent abnormal use if such use would engender a risk. In other cases the instructions must draw the user's attention to ways – which experience has shown might occur – in which the machinery should not be used. (Council of the European Union 1998) [emphasis my own]

What this means for the translator is that, if a manual describing the procedure for repairing a machine vital information is missing or incomprehensible and could result in death or injury or damage to the machine, the translator needs to ensure that the information is either completely reformulated, supplemented and made explicit or at the very least brought to the attention of the client. I once had to translate a maintenance manual for a die pressing machine used to manufacture body panels for motor cars. As such texts go this was fairly unproblematic except for one particular section

which, if I had translated as it stood, would have caused the user to lose lost an arm as a result of reaching in to the machine without first engaging the safety mechanisms. The manual was divided into various sections and stages which described the different maintenance tasks which could be carried out. In a previous section, the manual described how to disengage the power supply to the die punch and activate the safety barrier before reconnecting the power. In the next section, this information was left out although it was still essential. It's possible that the author assumed that it wouldn't be necessary to repeat the information a second time. Perhaps the writer simply forgot to include it. In any case, there are several compelling reasons as to why it was essential to include this information again:

- the nature of manuals means that people do not always read manuals from cover to cover or even in sequence,
- even if users had read the previous section, there is no guarantee that they would remember this information
- even if users realise that there is potentially something missing from the instructions, many people trust the instructions they are given and will assume that the information was left out for a reason, e.g. it does not apply in this particular instance

Ambiguous information can also necessitate the addition of information in a text. Pinchuck (1977:66) gives an example of ambiguity in the form of the following text on a sign:

> "Flying planes can be dangerous"

We can disambiguate this by making explicit certain information relating to the situation in which the utterance is intended to function or be used. Thus we could say:

> "Planes flying overhead can be dangerous".

While this may seem unacceptable because it is clearly adding text where there was none, if we take the view that a text is supposed to function in a particular situation, it therefore depends on that situation in order to have the correct meaning, relevance and importance. Thus, the text is part of the situation and so adding material or facts from the situation is perfectly acceptable because it is not so much adding but reallocating information. In

fact this is similar to what Kussmaul (1998:123) describes as "enlarging a scene". In this case, the scene would be the situation and circumstances in which the sign is used, e.g. on an airfield. He does urge caution in this respect because of the difficulty in definitively saying "how many and which kinds of elements can be added to a scene without getting into another scene that is no longer adequate as a translation". So there is a limit to how much information we can make explicit from the scene but unfortunately, there are no hard and fast rules.

Conversely, it is sometimes essential to remove information from a text. This can happen for several reasons: sometimes the information is legally inappropriate for a particular audience; the information does not apply to the target audience for technical reasons; or the information is meaningless for the target audience.

Another example comes from a translation I carried out for the manufacturer of a satellite television receiver. The document consisted of a series of frequently asked questions relating to the receiver, its configuration and its operation. The document stated that users could find details of the frequencies and polarities of various television channels in a range of German magazines, each of which was mentioned by name. At first glance this was a fairly innocuous sentence but in reality it was potentially problematic. Firstly, the magazines are not available outside Germany and even if they were, they would be in German and would be of no use to the reader. An alternative to listing the names of the magazines would be to do a little research and replace the names with the names of similar English-language magazines which also provided the same information. A little above and beyond the call of duty perhaps, but it is the type of thing some translators would do as a matter of course for a customer. This, however, is not without its problems either. With English being spoken in several countries, how do we know that the magazines we select will be available in places like Canada, the United States, Ireland, England, Australia and so on? The translation itself may end up in any number of these countries but the likelihood of a magazine being available in some or all of them is by no means assured. In this case, a more sensible solution would be to simply omit the names of the magazines and refer users to "your satellite listings magazine".

Theory in Technical Translation

Technical translation, like translation in general, has both benefited and suffered as a result of the work of translation theorists. In the past 40 or so

years a plethora of theories, models, approaches and ideas have been circulated seeking to explain, rationalise, analyse and describe the translation process. Technical translation has, however, been largely omitted from much of this work and is rarely dealt with explicitly. Rather, it is for technical translators themselves or those academics who feel strongly enough about this species of translating to try and see how the "mainstream" theories can be related to the practice of technical translation at the coal-face, as it were. More often than not, the results are disappointing. Not because technical translation is inherently unique and challenging (although it is) but because translation itself is as elusive and mercurial as ever and there seems to be a real collective mental-block preventing a clear understanding of what translators actually do. But what are our options? Do we want a theory of translation that is sufficiently flexible and general that it can be applied on at least some level to all types of translation but which will not provide any concrete answers or insights? Or do we want highly specialised, narrowly focussed theories that are so rigid and unyielding that they either exclude vast swathes of translation activity or come crashing down around our ears at the first sign of an anomalous translation situation?

The aim of this book is most certainly not to redress this situation. It would be a brave soul who would try to tackle the patchwork that is translation studies in this way. Instead, my aim here is to discuss various theories, models and approaches and try to relate them in some way to technical translation. Conscious of the need to avoid reinventing the wheel or worse still, further muddying the already murky waters of translation theory, there will be no lengthy or exhaustive discussions of the various theories. This is not necessary or indeed appropriate here. The aim here is, therefore, to provide some theoretical background so as to illustrate the nature of technical translation and to provide a better understanding of the environment within which it operates.

The Trouble with Translation Theory

> Good theory is based on information gained from practice. Good practice is based on carefully worked-out theory. The two are interdependent. (Larson 1991:1)

When trying to explain or situate technical translation within a theoretical framework it is often extremely difficult to know where to begin. This is made all the more problematic when we consider the shockingly diverse

range of approaches, models, rules and theories. This can be illustrated using a much quoted example from Savory (1957:49) who compiled the following list of "rules" of translation from a variety of "authoritative" sources on translation which state that a translation:

- must give the words of the original
- must give the ideas of the original
- should read like an original text
- should read like a translation
- should reflect the style of the original
- should possess the style of the original
- should read as a contemporary of the original
- read like a contemporary of the translation
- may add to or omit from the original
- may never add to or omit from the original

These rules are, without doubt, contradictory and at times paradoxical but they cannot be dismissed completely since each one will find legitimate application in a specific translation context. And this is, perhaps, where we should begin our attempts to situate technical translation in its own theoretical "space", if such a thing is possible. Looking at the professional context upon which this book has as its background, we can describe technical translation in very basic terms as a communicative process, or rather, a service provided on behalf of someone else for a particular purpose and within a particular situation and environment. In fact, the constraints and circumstances affecting technical translation are really quite similar to those affecting technical writing which will be discussed in the next chapter.

So how does this help us find a theoretical basis for technical translation? To answer this, we first need to look at the most prevalent theoretical approaches to translation to see where this communicative, reader-orientated, service approach fits in. We can broadly categorise the different theories according to the emphasis they place on the source text (ST) and/or target text (TT). It should be noted, however, that such is the situation with translation theory that there is frequently an overlap between these two categories; while a theorist may ostensibly be in favour of concentrating on the target text, there may be varying levels of source text-orientation within the theories. Far from being a ground for questioning the decision

to divide theories in this way, this fact serves to highlight not only the difficulty in developing a theory of translation but also in placing technical translation within a particular theoretical framework.

However, a useful starting point, particularly for technical translation, is the idea presented by Toury (1995:56-7) that translation is essentially affected by two major roles:

1. the translation is a text in a particular target language and culture
2. the translation constitutes a representation in one language of another text existing in another language and culture

He states that the basic choice which is made by a translator in deciding between the requirements of the two different roles represents an *initial norm.* This choice with regard to preferences for either role determines whether the translation can be regarded as *adequate* or *acceptable.* Thus he says:

> …whereas adherence to source norms determines a translation's **adequacy** as compared to the source text, subscription to norms originating in the target culture determines its **acceptability**. (Toury 1995:57) [emphasis in original]

We can see, therefore, that adequacy is typically regarded as a term associated with source text-orientated attitudes and approaches; acceptability on the other hand, by its very implication of a target audience focuses on the target text. Relating this to technical translation, we can say that, in general, the aim of a technical translation is to achieve a high level of acceptability, primarily because technical texts, particularly instructional texts, are intended to function first and foremost as a target language text.

Toury goes on to say that adequacy and acceptability can be manifest on both a macro and a micro-level. Thus, on a macro-level, the text may subscribe to target culture norms but on a micro-level, not all translation choices and strategies need to comply with the overall macro-level strategy. In effect, we can see that we can still produce a target-orientated text even if we have to resort on occasions to source-orientated strategies on a lower level. This is especially true of technical translation where some sentences can and need only be translated literally while other sentences or even paragraphs need much more radical translation strategies.

Larson (1987:69) provides some additional insight when she states that "the goal of most translators is to produce translations which are acceptable for the audiences for whom the translations are produced". This acceptability,

she continues, can only be attained "if the translator has a set of criteria which will guide him [sic] as he works and by which he can measure his final product". Crucially, she says that these criteria vary from project to project - or from translation to translation.

The notion that a technical translation must be acceptable to the client is echoed by Pinchuck (1977:205) who states that what both the client and translator want is a satisfactory translation which is achieved with a minimum expenditure of time and effort. It should be pointed out at this point, however, that Pinchuck does not make a distinction between acceptability and adequacy like Toury. While Pinchuck refers to the adequacy of translations, it is used in the same way as Toury uses acceptability - in other words, the translation is adequate for the needs of the target audience. Indeed, Pinchuck says an adequate translation is always a compromise between conflicting demands in that has a tendency to concede fidelity to the requirements of intelligibility and of speed. The extent of this, Pinchuck maintains, depends on the circumstances in which the translation is produced (1977:207).

In general, an inferior translation according to Pinchuck *(ibid.)* is "a work that is unreadable, or difficult to read, and possibly also inaccurate". Pinchuck goes on to say that in any case, an adequate translation will always be one that has been produced utilising just enough time and energy to ensure that the needs of the consumer are met. The translation should not be of a higher quality than that required by the client, particularly if this involves a higher cost. Having said that, the quality should not be lower if it means that the reader will have to spend a lot of time and energy on deciphering it. The technical translator, according to Pinchuck, aims to achieve adequacy and not perfection in translation. Realistically, this should be the goal for any translator or writer because we can never assume that a communicative act will run perfectly smoothly and that nothing will be lost.

Source-Orientated Approaches to Translation

Equivalence

The idea of equivalence forms the basis of many theories of translation and by implication, definitions of translation quality. But the term "equivalence", however, is fraught with difficulties. Indeed Newmark (1993:75) claims that "the cerebration and the brain racking about translation equivalence

goes on forever". He goes on to say that translation equivalence cannot be defined and as such, there are only *degrees* of equivalence *(ibid).*

While there are numerous definitions and types of equivalence, they all rely on one thing: a link or bond of some sort between the source text and the target text. As Catford (1965:49) explains "the TL text must be relatable to at least some of the situational features to which the SL text is relatable". According to Catford, a translation does not necessarily have to consist of target language elements with the same meaning as the source language elements but ones which have the greatest possible amount of overlap across the different levels of equivalence.

This is echoed by Halverson (1997:207) who provides a slightly broader definition of equivalence and says it is a relationship which exists between two entities and it is described as one of likeness / sameness / similarity / equality in terms of any number of potential qualities. Equivalence theories regard translation as an attempt to reproduce the source text as closely as possible. But because there are numerous reasons why they can never be perfectly equivalent on all levels, numerous of types equivalence have been defined such as formal and dynamic equivalence (Nida 1964), denotative, connotative, pragmatic, textual and formal aesthetic equivalence (Koller 1979).

Levels of Equivalence

We see that equivalence between the source and target texts is not absolute and total but it can operate on a number of different levels. There are a number of systems which have been put forward to examine the levels of equivalence. One of the most enduring is the scheme proposed by Koller (1979:188-189). According to Koller, equivalence can occur on the following levels:

1. *Denotational Meaning*, namely the object or concept being referred to. In texts such as user guides, equivalence on this level is relatively easy to achieve because we need to keep referring back to the product the reader is learning to use.

2. *Connotational Meaning*, which is, according to Koller divided into language level, sociolect, dialect, medium, style, frequency, domain, value and emotional tone. Equivalence on this level can sometimes prove problematic where, for example, the enthusiastic and informal tone adopted by the source text may be culturally unacceptable in the target language.

3. *Textual Norms*, which are typical language features of texts like patents, legal documents, business letters etc. Textual norms may include the use of the second person to address the readers of user guides or the use of passive constructions to describe experiments in scholarly journals.

4. *Pragmatic Meaning*, which includes reader expectations

Komissarov (1977) also proposes a series of levels on which translation equivalence can occur. In the context of what Fawcett (1997:60) refers to as "sharp and fuzzy equivalence" Komissarov's levels all build upon each other with the lowest level of equivalence being retained and built on by the next higher level of equivalence and so on. These levels are as follows:

1. Equivalence on the level of the general meaning or message. Aside from the general communicative intention, there is no tangible equivalence on the basis of situational, grammatical or lexical factors.

2. In addition to the preceding level, this type of equivalence identifies a particular situation in both texts.

3. Building on the preceding levels, equivalence on this level maintains factors or elements which describe the situation established in level 2.

4. As well as all of the information presented as a result of the preceding levels of equivalence, this level establishes semantic and syntactic equivalence in that the majority of the source text words have corresponding target language lexical items and the syntax is transformationally related.

5. This level of equivalence displays a close parallelism between the source and target texts on all levels of language.

Types of Equivalence

Perhaps the most well known types of equivalence are formal and dynamic equivalence posited by Nida (1964). Formal equivalence is concerned with the message in terms of its form and content. With this type of equivalence the message in the target language should match the different elements in the source language as closely as possible, be they lexical, syntactic, stylistic, phonological or orthographic. According to Catford, a formal correspondent (or equivalent) is any target language category (unit, class, structure, element of structure, etc.) which can be said to occupy, as nearly as possible, the 'same' place in the 'economy' of' the target language as the given source language category occupies in the source language (Catford 1965:27).

Dynamic equivalence on the other hand is based on the notion that the target text should have the same effect on its audience as the source text had on its own audience. With this type of equivalence the emphasis is not so much on finding a target language match for a source language message but rather on creating the same relationship between the target audience and the message as that which existed between the source language audience and the message (Nida 1964:159). By using this type of equivalence it is hoped to produce a target text which is natural and idiomatic and which focuses on the target language culture. According to this definition of equivalence, a successful translation needs to capture the sense of the source text and not just the words. As such it can only be regarded as a successful piece of communication if the message is successfully transmitted to the target audience. One example, proposed by Nida, of how this can be achieved in practice would be to change the sequence of sentences where they do not match the real-time chronology of actions. This particular strategy is quite useful, especially in the translation of instructions.

On the surface, dynamic equivalence seems useful in that it emphasises the importance of meaning, and more specifically, of equivalent response. It could be argued that this is essential in the translation of instructional texts, training materials, tender documents etc. but the usefulness of this concept is limited as it is primarily concerned with cultural phenomena which, as a rule, do not occur in technical texts. I say *as a rule* because technical texts can, and do on occasion, contain certain items which Lee-Jahnke (1998:82) describes as *socio-culturally specific*. Such items may include time and date conventions, units of measure, laws, government departments and authorities etc.

While legal terms, such as those that appear in product documentation or patents, are potentially problematic, from a practical point of view it can be argued that they are easier to deal with. If for example, the target audience needs to *know* precise details of the law, we give the name of the law in the source language accompanied by the official translation or a paraphrase in the target language. The reason for including the source language name here is to ensure that if the reader needs more information on the law or needs to seek advice from a lawyer in the source language country, they will be able to refer to the law using its correct name. There is, after all, little point in referring to something by a name which nobody recognises. On the other band, if the precise term is not important, we simply give a description or generic term. An example of this would be a user guide for a refrigerator which tells users to dispose of the packaging in accordance with the German laws on disposal of household waste. It makes little sense for someone living outside Germany to dispose of waste in accordance with

German laws. In this case, we simply say that the packaging should be disposed of in accordance with the relevant waste disposal laws (unless of course we are certain of' the intended target country and want to mention specific equivalent laws but sometimes it's best to keep these things vague!).

Nida makes the point, however, that eliciting the same response from two different groups of people can be difficult, particularly when we consider that no two people from the same language group will understand words in exactly the same way (1969:4). This sentiment is also expressed by Steiner (1975:28). What we are left with, therefore, is an approach which is theoretically quite desirable but often regarded as excessively vague, difficult to implement and imprecise in practice.

Other types of equivalence include *referential equivalence* whereby the equivalence centres on the ST and TT words referring to the same extratextual entities in the "real world". *Connotative equivalence* is used to describe the relationship between ST and TT words which result in the same or similar associations or connotations in the minds of the reader. *Text-normative equivalence* involves source language and target language words being used in the same or similar contexts in their respective languages.

Technical Translation and Equivalence

Source-based approaches, as exemplified by the various types and levels of equivalence briefly mentioned above, represent a problematic foundation upon which to base technical translation. While referential equivalence can ensure that a translation accurately conveys the intended information, connotative equivalence can help avoid the introduction of inappropriate register or terms and textual equivalence can benefit the flow of information and cohesion of texts, the fact that source-based approaches do not consider the full communicative situation in which technical texts are translated and used poses significant problems for the technical translator. Since technical translation is a communicative service aimed at providing information to a new audience, the concentration on the source text and not on those involved in the communication means that a crucial part of the translational situation is simply not considered. If we do not consider the purpose of the communication, it will be difficult, if not impossible, to tell whether it was successful.

Linguistic approaches - of which Nida's theory is one, its sociological dimension notwithstanding - presuppose some form of equivalence between the source text and the target text. This is fine as long as we are solely interested in the text and we are content that the target text, as a

reflection of the source text, is an end in itself. But if we are interested in the communicative value of a translation and what people do with texts (see Robinson 2003:142), equivalence-based theories have difficulty in accounting for the changes, alterations, additions and omissions etc. which are needed in professional translation projects. Koller (1995:196) acknowledges this when he talks about the “contradictory and scarcely reconcilable linguistic-textual and extra-linguistic factors and conditions”.

Source-based approaches also fail to take into account the fact that translations, once they have been “released into the wild”, so to speak, become subject to the norms, standards and requirements of contemporary texts originally produced in the target language. In other words, the translation is no longer regarded by the target audience as a translation and instead is measured against other target language texts. Rather than providing us with a means of producing independent and autonomous target language texts, equivalence, because of its need to maintain a close link between source and target texts, provides us with texts that can only be evaluated on the basis of a source text which the target audience will usually not know about. Apart from this, if the target audience was in a position to compare the translation with the source text, they would be unlikely to need a translation in the first place.

Problems also arise from the fact that the various typologies of equivalence rarely provide any real guidance as to how we should go about actually translating texts. While equivalence typologies such as Komissarov’s above are useful in highlighting the different levels of equivalence which may be achieved by a translator, when it comes to the actual process of translating they are difficult to implement because they do not specify which type of equivalence could or should be used under which circumstances. For instance, in the case of a user guide, should the translator strive for denotational equivalence alone or denotational and textual equivalence? As Fawcett (1997:62) says “it is only a little more helpful than the old translation adage ‘as literal as possible, as free as necessary”. Some source-based approaches, do in fact, provide some form of prescriptive rules to be applied during the translation process, for example Schveitser (1987). Though interesting and enlightening, they are not particularly useful for practical applications because they are generally too cumbersome and numerous for one person to remember and implement. In any case, in schemes such as Schveitser’s which has some 55 rules, only a small proportion of the rules will apply in any one translation job. The time spent by a translator trying to decide which of these rules applies to the various parts of the text would be better spent actually reading the text or researching parallel texts.

Functionalism

Unfortunately for the professional translator, the categories, levels and classifications of equivalence described above, while helping us to pick through a translation to see how it ticks, do not really help with the actual process of translation.

In an attempt to escape the restrictive and often limited approaches to translation based on theories of equivalence, translation theorists such as Reiss (1971) and House (1981) changed the focus from being entirely source-based to include some aspects of the target text. To be precise, their attention centred on the function of the target text. Such an approach moves away from the bottom-up linguistic approaches of equivalence-based theories and instead involves pragmatic and situational aspects of the translation process. This is indeed an improvement in that it goes at least some way towards acknowledging the fact that texts are written and translated for a reason. However, functionalist based theories do not entirely forsake the source text. Rather they are a hybrid approach which considers both the source and the target texts.

Nevertheless, functionalism as a general ideology based on extralinguistic, pragmatic and communicative factors of translation is nothing new. Even in 1964, Nida's notion of dynamic equivalence called for the reproduction of the effect (or function) of the source text in the target text through equivalence of extralinguistic communicative effect. This was already hinting at functionalism as we later came to know it.

In 1971, Reiss included the element of text function in her model of translation criticism. While Reiss' work is often regarded as highly conservative and dated, it did mark a turning point in the way scholars looked at translation, particularly in Germany. Her model, while being overwhelmingly equivalence-based, also includes the functional relationship between the source and target texts. According to Reiss, the ideal translation is one where optimum equivalence is achieved as regards the conceptual content, linguistic form and communicative function. This essentially means that, taking into account the linguistic and situational context, the linguistic and stylistic factors as well as the author's intentions (even though this in itself is a persistent source of debate), the target text should have the same "value" as the source text. One of the problems, however, with this approach to the function of translations is that it cannot deal with instances of translation where the function of the target text is different to that of the source. To circumvent this problem, Reiss defines such instances as *"Übertragungen"* (1971:105) or transfers. Thus, translations where the function changes are not, she maintains, 'real' translations, but rather adaptations.

House (1981) also adopts a functionalist approach and states that it is "undeniably true that a translation should produce equivalent responses" (1981:9) and it is on this basis that she maintains that the ultimate goal of a translation is to achieve the same function in the target text as that in the source text. She defines two types of translation: *covert* and *overt*. A covert translation is one where the text function is preserved and the reader is not aware that the text is a translation. An overt translation, on the other hand, does not maintain the text function of the original and the readers are somehow aware that it is a translation and not the original language text. In order to determine whether functional equivalence has been achieved, Houses proposes that the source text be analysed first so that the target text can be compared against it.

Like Reiss, this approach escapes the restrictive, purely linguistic criteria of traditional equivalence-based models by including certain pragmatic and extralinguistic factors but they both fail to take into instances where it is either not always possible or not desirable to maintain the same function in both texts (cf. Kade 1977:33; Nord 1997:9) to accommodate audience expectations for the text genre (cf. Gerzymisch-Arbogast 1993:33-34). A prime example of this comes in the analysis of a tourist brochure by House (1981:118ff). House criticises the translator for not preserving the apparent flattery provided by the German text's lack of explanation for certain culturally-specific items. She regards this, not so much as an error, but more as proof that the translation was not really a translation, but rather an adaptation. Gutt (1991:50) asks the question of what happens if flattery is not socially acceptable in the target culture. Can the translation really be criticised for conforming to the social norms and customs of the target audience? After all, preserving individual functions within a translation may ultimately make the translation as a whole functionally non-equivalent. It could be argued that purposely changing the function in the target text is actually an important way of hiding any clues as to the text's true origins in the source language, something which is central to the notion of a covert translation.

According to House texts where the function is not maintained are *overt* (1977:194) and not real translations because they draw attention to the fact that they are translations. Reiss claims that a target text where the function of the source text is not maintained is a *transfer* and dismisses them as something other than "real" translations. So for instance, if a German user guide is intended to *instruct* while an English user guide is intended to *explain*, a translation between these two languages will not be a translation but rather an adaptation, even though the target text can be regarded as a translation in every other sense.

What is more – as is the case with traditional equivalence-based approaches – any attempt at translation based on this approach is constrained by the need to define the target text in terms of how closely it reflects the source text on a variety of levels. This is problematic in that the target text is not regarded as an independent, autonomous text – and this is how translations are regarded by the target audience, at least in a professional context. Since Reiss's model, for example, focuses on the linguistic means the source language uses to express text function, the wording of the source text becomes the yardstick for judging the appropriateness of the linguistic means employed in the target text (Lauscher 2000:155). But because text function is expressed using different linguistic means or writing strategies in different languages, reflecting the source text linguistic features will result in unusual and possibly unacceptable target texts.

There are other problems with this type of approach, however. Nord (1991:23) argues that the function of the target text cannot be "arrived at automatically from an analysis of the source text". Instead, it needs to be defined pragmatically by the purpose of the intercultural communication. Fawcett (1997:107) maintains that there is no need to link text function and translation strategy. He maintains that just because it is possible to identify the function of a text, there is no "logical or translation-scientific imperative" on a translator to shape and govern translation decisions on the basis of the function. Besides, if technical texts are translated with a specific communicative situation in mind, surely this situation should form an integral part of the translation process and define the end result. In any case, using the source text as the sole means for determining the function of the source text and subsequently the target text is a risky business. This is primarily because such an approach presupposes a uniformity of skill and care on the part of the original author in the way the text is written (thereby encoding the text function within it) which is impossible to quantify or achieve.

Technical translators also encounter the problem that it is not always apparent from the text what type of text it is, let alone what the actual purpose is. This can be for a variety of reasons: the text sent to the translator may be pre-translated or tagged using a translation memory (TM) tool and the visual clues indicating text type may be missing making the linguistic clues more difficult to spot; or web pages may be cut and pasted into a word processor file. In other cases, the file to be translated may be designed to serve as the source document for a single-source multi-channel publishing system where the same text is used to provide content for a range of texts such as manuals, advertising materials, newsletters, technical documentation, web pages, presentations, etc. In such cases, the translator cannot

possibly know what the ultimate function of the translated text will be unless some form of extratextual information is available.

Target-Orientated Approaches to Translation

Toury (1995:26) in his work *Descriptive Translation Studies and Beyond* puts forward the notion that the position and function of translations "are determined first and foremost by considerations originating in the culture which hosts them". Toury regards translations as "facts of target cultures" and in his discussion of traditional methods of examining texts he mentions the fact that they "were primarily concerned with the source text and with its inviolable 'sanctity" whereby target text factors "while never totally ignored, often counted as subsidiary especially those which would not fall within linguistics of any kind" (Toury 1995:24). Here we see a shift in emphasis away from the source-orientated treatment of translations and towards the treatment of translations as autonomous, independently functioning target language texts. And this is precisely how the translations are generally viewed by readers.

The following paragraphs deal with approaches to translation which do not depend on the supremacy of the source text. In other words, theories which acknowledge the importance of the target text in the translation process.

Relevance

Gutt (1991:22) stresses that translation should be approached from a communicative point of view rather than any other, for instance, a theory of translation. Basing his relevance theory on work by Sperber and Wilson (1986), Gutt proceeds from the notion that the key to human communication is our ability to draw inferences from people's behaviour, be it verbal or non-verbal. Thus, what is said (or written) provides, within the context of relevance theory, the stimulus from which the recipient can infer what the sender means. This is referred to as the informative intention.

Breaking the communicative process down into steps, Gutt maintains that utterances are firstly decoded and, on the basis of their linguistic properties, are assigned to meanings or things they represent. This is similar to de Saussure's notion of *signifiant* and *signifié* or *aliquid statt pro aliquo* [something stands for something else] (Linke *et al.* 1994:18, 30). Thus, these representations are referred to as semantic representations and they refer to the mental representations which are essentially the output of the

mind's "language module[5]" (1991:24). He continues to say that these semantic representations are assumption schemas which need to be processed in order to become truth-conditional and to have a propositional form. In other words, out of all of the possible meanings or representations associated with an utterance, the brain must select what is in fact feasible, likely or credible within the particular context of the utterance. Indeed, it is the very notion of context which helps, according to Gutt, to explain the fact that while linguistic expressions do, in fact, have a meaning, it is not necessarily the same as that conveyed by the expression at a particular time and place (1991:25).

Gutt regards context as a psychological concept which, in contrast to common definitions, does not refer "to some part of the external environment of the communicative partners, be it the text preceding or following an utterance, situational circumstances, cultural factors etc., it rather refers to part of their 'assumptions about the world' or cognitive environment" (1991:25). Cognitive environment does not exclude the various external factors but rather includes them along with information they provide and its "mental availability for the interpretation process" (*ibid.*).

But we find ourselves asking the question of how readers select the correct assumptions and meanings as intended by the sender, or author. This is indeed essential for communicative success. Gutt asserts that communication is governed by a desire to "optimise resources". In other words, readers want to gain as much information with as little expenditure of resources as possible. This means that meaning and understanding will always take the path of least resistance, and readers will start processing information using those contextual assumptions which are most accessible to them. He goes on to say that "[a translation] should be expressed in such a manner that it yields the intended interpretation without putting the audience to unnecessary processing effort" (1991:101-102). Relevance theory might well require us to build in certain contextual clues for a reader to fully and correctly understand a text. This modification and addition of information in the text brings us towards the notion of interlingual technical writing mentioned elsewhere.

---

[5] This notion of a language module is presumably made for the sake of convenience. Although we can associate language skills with certain parts of the brain – usually in the left hemisphere (Eysenck 2000:79; Kalat 2004:441) – there is no specific "language module" to speak of in the human cognitive system; language is regulated by a range of processes and functions throughout the cognitive system.

Thus, the assumptions and information which allow the reader to understand the communicative intent must be easily accessible to the reader. Of course, when applied to the text we can deduce that this information can be composed of both contextual information in the form of knowledge of the world and also information presented in the text itself. Gutt says that information obtained through perception is generally given more importance than information based on inference and so it gives the translator greater flexibility to ensure that the first interpretation of the utterance which the reader arrives at is the correct one as intended by the sender. And perhaps equally as importantly, particularly in the case of communicative texts and manuals to ensure that the effort involved in obtaining the meaning or contextual effects is justified (1991:30-31). So, the responsibility ultimately falls to the translator to decide how best to achieve this to ensure that the contextual effects are "adequate to the occasion in a way the speaker could have foreseen" (1991:31). Such a notion of target audience expectations is similar to assertions by Pinchuck (1977) who claims that elements in texts provide "triggers" for readers to allow them to anticipate what is to follow next. Indeed, Gutt states this even more explicitly when he says "introductory words would guide the hearer in searching his memory for the intended referent and hence considerably ease his processing load" (1991:33). Similarly, Gerzymisch-Arbogast (1993) states that information must be presented in varying amounts of given and new information - relative to what the author perceives the audience to already know - in order to achieve the appropriate register and flow of information.

Hönig (1997), however, levels the criticism at relevance theory that it fails to take into account precisely what readers regard as a good or bad translation, only what they regard as relevant. It is indeed a sound point in that we do need to take into account what readers regard as good and bad. It would be conceivable to determine what readers regard as good or bad texts either from empirical studies involving reader reactions or usability studies or to analyse a broad range of style guides relating to the text type in question.

Relevance theory does provide compelling support for taking the cognitive abilities of readers into account to improve translation but Gutt's opinions regarding what does and does not constitute a translation is quite problematic for technical translation particularly with regard to interpretative and descriptive translations.

At the heart of relevance theory is Gutt's distinction between interpretive and descriptive language use and it is this distinction which is most relevant in terms of technical translation. Interpretive language use, according to Gutt, refers to an utterance which is intended to represent what

someone else thinks or says. In the context of translation, an interpretive translation is one which has some form of link or relationship with the source text. Descriptive language use, on the other hand, is intended to be an accurate representation of what the writer or speaker believes to be true. In the context of translation, a descriptive translation is intended to function as an independent and autonomous target language text whose readers will not be aware of the existence of a source text.

According to Gutt, only an interpretive translation is a "real" translation. A descriptive translation, he claims, cannot be regarded as a true translation because a translation only achieves relevance by communicating effectively AND by standing in for the original text. While Gutt maintains that a translation must share at least some link with the source text, the notion that a text which can survive on its own and which is undetectable as a translation cannot be regarded as a true translation is difficult to reconcile with the reality of translation.

We could argue that a descriptive translation mimics the production of an original text but it is still a translation albeit one that could have been achieved using very different means. Gutt's assertion that only interpretive translations are true translations seems to contradict the communicative orientation of his theory if translation is a communicative act, then the translation should do the same as a standard text or communicative act, i.e. convey information in order to affect a change in behaviour, attitude or knowledge of the receiver (cf. Gommlich 1993:177-178).

He does try to clarify interpretive use by making two sub-categories: direct and indirect translation. Almost as if he realised that the existing categories of descriptive and interpretive would effectively eliminate translations whose function was to blend seamlessly into the target language landscape, he says that indirect translation allows the translator to "elaborate or summarize" (Gutt 1991:122) in order to communicate as clearly as possible rather than "give the receptor language audience access to the authentic meaning of the original" (Gutt 1991:177). This certainly bears a remarkable similarity to the aims, factors and features involved in a descriptive translation.

Gutt claims that if a translator aims to produce a translation that can stand alone as a target language text using whatever means necessary, then it is not a "real" translation but rather an "adaptation", in principle this might be true but there is a problem with it insofar as technical translation is concerned. Some texts require more work than others. For example, a technical specification or description of a process will require less work than a user guide or technical brochure. So while the technical specification or

brochure are intended to serve as freestanding original texts, they cannot be regarded as anything but adaptations because there has been no "adaptation" beyond that which is necessary to ensure compliance with target language norms, requirements and conventions.

On the other hand, a user guide which needs to be modified in a number of ways, e.g. register, tenor, cognitive suitability, legal compliance etc. can only be regarded as an adaptation. Here the ST has been used as the basis for producing an original target language text. The problem is that most people would argue that the user guide is still a translation and not an adaptation. Perhaps the problem is really one of scales. If there are huge amounts of adaptation, omission and addition then it becomes an adaptation. What we are then left with is the problem of how much is too much?

Nevertheless, despite the confusing approach to what does and does not constitute a "real" translation, relevance theory does provide some useful insights into technical translation:

- a communicative approach which concentrates on the needs and expectations of the target audience
- the minimax principle which, put simply, states that people do not want to spend more time and effort than is absolutely necessary in order to retrieve information from a text; this is especially true for technical texts such as user guides.

Skopos Theory

Skopos theory is linked with the functionalist approach to translation but differs fundamentally in that where functionalist approaches such as Reiss's (1971) and House's (1981) maintain that the function of the target must be the same as the original, Skopos theory recognises that this is not always practical or desirable. According to Vermeer, the methods and strategies used to produce a translation are determined by the intended purpose of the target text (Vermeer 1978:100). Unlike equivalence-based theories where the source text and its effect on the source language audience or even the function attributed to it by the author determine the translation, Skopos theory holds that the prospective function or Skopos of the TT as determined by the initiator (the person who initiates the translation process, i.e. the client) and the translator. As such, the Skopos is determined by the initiator / customer, their view of the target audience along with the situational and cultural background. Skopos theory states "that one must translate consciously and consistently, in accordance with some principle

respecting the target text. The theory does not state what the principle is: this must be decided separately in each specific case" (Vermeer 1989:182). In conjunction with this top-level rule are the supplementary general rules of coherence and fidelity.

According to Nord (1997), the coherence rule maintains that the target text should be sufficiently coherent in order for the reader to comprehend it. Essentially, this rule requires that the TT fulfils the basic requirements for any text written in that language. The fidelity rule is less perspicuous in that it requires that there be some sort of relationship between the ST and TT once the Skopos and coherence rules have been satisfied.

Unlike traditional functionalist theories such as that put forward by House (1981) which focus on the pragmatic aspects of the translation act but where the source and target texts have the same function, Skopos theory acknowledges equivalence of function as just one of many possible Skopoi of a translation. The reason for this is that there are numerous situations and circumstances where a translation needs to be produced. As such, the situational environment of the translation process determines the Skopos as does the text receiver. Thus the Skopos of the source text and the target text may be different because of the needs of the two audiences by virtue of the fact that they belong to two different social and linguistic realities. Cases where the function stays the same are referred to by Reiss & Vermeer (1991:45) as *Funktionskonstanz* (unchanged function); cases where it changes are referred to as *Funktionsänderung* (changed function). Kade (1977:33) also makes this distinction, referring to the fact that the function of the translation does not have to be the same as that of the source text, he defines equivalent translation as that which retains the communicative function of the source text and heterovalent translation as that which involves a reworking of the content and where the target text takes on a different function to that of the original. Nord (1997:9) also recognises this fact when she states that we do not arrive at the function of the target text from an analysis of the source text but rather from a pragmatic analysis of the purpose of the communicative act. Referring back to Skopos theory, she maintains that functional equivalence (the source and target texts having the same function) is not the "normal" Skopos of a translation but only one of a number of potential Skopoi and one in which a value of zero is assigned to the "change of functions" factor (1991:23).

First and foremost, however, the Skopos of the translation must be formalised and clearly set out before the translator can actually start work. The process of defining the Skopos of a translation is included in what is called the translation brief (Vermeer 1989, Kussmaul 1995 and Nord 1997). We can look at the translation brief as a form of project specification which sets

out the requirements for the service / product to be provided. Ideally, such a brief would be quite specific about the intended function of the translation, the target audience, the time, place and medium as well as purpose. The problem with this concept is that the client who initiates a translation is rarely a language professional and usually has no specialised linguistic knowledge. Many clients have no interest whatsoever in the "mechanics" of the translation process and may even regard such information as the responsibility of the translator.

Here the translator needs to function much like an architect who discusses a building project with an ordinary, lay customer. The customer comes to the architect with an idea for a house and the architect advises what is and is not possible within the constraints of physics, materials science and building regulations. Having established what the customer wants, the architect then decides how to design and build the house and what materials will be used. The customer will not specify which structural materials will be used but may specify cosmetic materials like tiles, glass, doors, balustrades etc. This distinction is equivalent to the translator looking after translation strategies and linguistic matters and the customer specifying such things as style, terminology or general audience. In both cases, the customer specifies the objective and the expert (in our case the translator) decides how best to achieve that. Thus, the Skopos is not a random or accidental occurrence – it is specific to a particular constellation of factors relating to situation, purpose, requirements etc.

In defining translation as the production "of a text in a target setting for a target purpose and target addressees in target circumstances", Vermeer (1987a:29) presents the view that the target text is the foremost concern in translation acts. The source text, through its noticeable absence from the above definition, is of lesser importance here than in equivalence-based theories of translation discussed earlier. Indeed, Vermeer refers to the source text as an offer of information or *Informationsangebot* which is then turned, either wholly or partly, into an offer of information for the target audience (Vermeer 1982). Nord says of this offer of information that

> …any receiver (among them, the translator) chooses the items they regard as interesting, useful or adequate to the desired purposes. In translation, the chosen informational items are then transferred to the target culture… (1997:25-6)

It can even be argued that the source text merely represents the "raw materials" (Vermeer 1987b:541) for the translator in the production of a

target language text. Vermeer (1982) discusses the concept of *Informationsangebot* so that each of the countless potential receivers can select the relevant or interesting information from the text. As technical translators, our job is to ensure that this information is as easy to find select and assimilate as possible. This is quite similar to Gutt's relevance theory and the notion of usability engineering in texts which is discussed in Chapter 4. Rather than leaving the choice of which information to select completely to the reader, we can foreground information and make it more likely to be selected by readers.

A similar sentiment was expressed by Pinchuck (1977:215) several years previously when he said that "a text will normally contain more information than is needed". He goes on to say that not all of the information in a text is of equal value (1977:220). This is indeed fascinating as it signifies a break from traditional doctrine that all of the information in a text is essential. Thus we can, perhaps, on this basis justify the elimination or omission of information which is "unnecessary", irrelevant or unhelpful in relation to the communicative act. Drawing from this notion of redundancy, Pinchuck formulates more criteria for assessing the effectiveness or adequacy of a translation, namely the excess, deficiency or appropriateness of redundancy in a text. Claiming that different languages may adopt more formal styles than others, Pinchuck states that German, for example, is more given to adopting a more formal tone than English. As a result, the level of abstraction may also be higher in German than in English. This will inevitably result in the need for greater explicitation and redundancy in an English translation, for example by referring to the product specifically, repeating information to make the text more cohesive or to reinforce a certain point etc.

Loyalty to the Source Text

Unless the translator is told otherwise, the text type of the translation will be the same as the source text. To use Vermeer's example (1989:183), it is safe to assume that a scientific article on astronomy will be translated as such in the target language. This might seem to contradict the claim made elsewhere that consistency of function is not the normal Skopos of a translation but it is not the case. The function of a given text type in one language might be different to a greater or lesser degree in another. So while the purpose of the text might stay the same (i.e. the astronomy article stays an astronomy article), the function of the text, e.g. to entertain and educate or educate and provoke, may need to change depending on the culturally or text specific conventions in the target language. This is what Nord regards as a "conventional" assignment (1997:31). But in order to do this, it is

necessary to decide on various strategies, e.g. word-for-word, transposition, calquing, modulation etc.

Nord refers to the difficulty in speaking of a single, definitive text, instead claiming that the selectability of information presented in the *Informationsangebot* means that there are, in theory, multiple "texts" and multiple meanings and interpretations. Similarly, there are multiple situations in which the information could be used. Nord sums this up by saying that "a 'text' is as many texts as there are receivers" (1997:31). Hatim (2001:76) describes this as a text's "plurality of meaning".

With this potential for different interpretations and situations, translators need to ensure that the target text will at the very least be meaningful for the intended receivers and that it will make sense in the particular communicative situation and culture. Only if the text makes sense in the receiver's situation will the translation be regarded as successful. For this reason, Vermeer introduces what he calls the *coherence rule* which states that a translation must be coherent with the receivers' situation or that it should be part of the receivers' situation (Reiss and Vermeer 1984:113). This is defined as intratextual coherence and is one of the two basic types of "conceptual connectivity" central to Skopos theory (Hatim 2001:76).

This rule firmly places the emphasis on the target text and its audience. However, because a translation is an offer of information based upon a previous offer of information, there has to be at least some relationship between the source text and the target text lest there are criticisms that the translation is not a translation but rather some abstract form of adaptation (Nord 1997:32). Vermeer calls this intertextual coherence or fidelity. The form taken by intertextual coherence, however, depends both on the Skopos and the translator's interpretation of the source text (Nord 1997:32). Subordinate to the intratextual rule, the fidelity rule can take form ranging from an extremely faithful rendering of the source text or a loose rendering.

Nord (1997:125) introduces the notion of "function plus loyalty" to Skopos theory. This she defines as a bilateral commitment on the part of the translator to both source and target sides of the translational interaction. She stresses that it is not the same as fidelity or faithfulness which is associated with texts. Rather, loyalty is a relationship between people. It is the translator's responsibility to provide readers with what they expect of a translation (i.e., literal, word-for-word, free, readable). It also requires the translator to convey the author's intentions as regards the communicative interaction.

It is important to note that this is an improvement on the notion that the author's intentions are evident from what the author has actually written because this presupposes consistent and uniformly competent authors, something which is impossible to gauge in the real world. However, Vermeer himself expressed concerns in 1996 at the inclusion of the loyalty rule as it represented a social and subjective criterion which damaged the applicability and reliability of his general theory.

The Applicability of Skopos Theory to Technical Translation

There are claims that Skopos theory produces "mercenary" translators who are, in effect, only interested in pleasing their paymasters (Pym 1996:338). This is untrue for a number of reasons. The "dogma" referred to by Newmark (1990:105) simply does not exist. Also, Nord (1997:118-119) maintains that translators are only mercenaries when they do not use their own judgement as language professionals to argue for doing something a particular way.

It can also be argued that Skopos promotes or even depends on adaptation and as such cannot really be regarded as translation. This is not, strictly speaking, true. Granted, in some rather extreme cases there may be several quite drastic modifications needed in order to ensure that a translation fulfils its Skopos and meets the needs of the addressees/users, but in the vast majority of cases (at least in my experience as a professional technical translator), there is only rarely a need for wholesale modification or rewriting to the extent that the translation has only the most vague link with the original text.

This is precisely why Nord (1997:125) introduces the notion of "function plus loyalty" to Skopos theory mentioned above. This is intended to ensure that even though there may be significant differences between the source and target texts, the needs of the key stakeholders are met.

The fact that Skopos theory states translations must be produced "in accordance with some principle respecting the target text" (Vermeer 1989: 182) which must be decided separately in the translation brief for each specific case could be regarded as something of a shortcoming of Skopos theory. The reason being that it introduces an element of subjectivity into the process and is not always repeatable. However, this is not unique to Skopos theory and is inherent to many other approaches to translation. Consequently, this has to remain a necessary confound in the equation.

Conclusions

There are numerous translation strategies available to translators such as modification, transposition, paraphrasing, literal translation, calquing etc. All of these strategies are essential at various times when dealing with various different texts. The problem facing us as translators is that depending on which theory of translation we subscribe to, one or more of these strategies may be prohibited, or at best, inconsistent with the spirit of the particular theory. Indeed, these strategies are so different in terms of their effects that it seems impossible that they could be comfortably incorporated within a single "theory". For example, the more formal linguistic approaches might have difficulties in permitting the use of paraphrasing, primarily because they do not regard paraphrases as "real" translations. With approaches based on formal equivalence at its most conservative, anything other than word-for-word translation would be unacceptable. Similarly, it is conceivable that if we rigorously pursue the goal of dynamic equivalence as described by Nida, we would be prevented from using calques or introducing loan-words. Yet as translators, we use virtually all of these strategies at various stages, sometimes even within a single project, text or paragraph. And this seems to be part of the problem with much of the literature on translation, namely that a translator can have one overall aim for the translation, for example a target-orientated translation, but still use approaches and strategies ranging from literal translation to radical paraphrasing, addition and omission without compromising the communicative purpose of the translation process.

The idea that no one theory of translation, e.g. free, literal, formal, dynamic etc., can completely explain technical translation is borne out by Skopos theory. Nord (1997:29) explains that the Skopos of a particular translation project might require either "a free or a faithful translation, or anything in between... depending on the purpose of the translation". Indeed, within a given text, we may need to switch frequently between literal and free translation for individual phrases, sentences or paragraphs.

It would seem, then, that the easiest way of maintaining all of the tools and strategies in the translator's repertoire is to adopt a communicative approach to translation like that proposed by Gutt but without the ambiguity as to constitutes a translation. After all, texts are written to communicate information and translations are no exception. Thus, Skopos theory would seem to provide a more flexible framework which ensures that the strategies and techniques of translation can be used legitimately wherever the translator deems it to be appropriate. We simply select the appropriate strategies

depending on the Skopos for the project and the text. As Nord maintains, "the translation purpose justifies the translation procedures" (1997:124).

This is not without its problems, however, particularly as translators are left to make decisions which can be regarded as, at best, reasoned but *ad hoc*, or at worst, subjective. But this problem is inherent to all theoretical approaches to translation and is not unique to Skopos theory. However, it is my view that Skopos theory is the only approach that truly acknowledges the professional reality of translating and the demands, expectations and obligations of translators.

To conclude this chapter, it is worth restating some of the more important areas that need to be considered when translating technical texts:

1. we need to concentrate on the needs of the target audience this is who the translation is produced for and these are the judges of whether a translation is actually good or not;
2. we need to understand what it is the target audience needs and wants;
3. we need to understand how technical communication works in the target language if we are to produce independent. autonomous texts that can "compete" with other texts produced in that language;
4. we need to remember that it is necessary to add, change or remove information as part of the translation process in order to achieve effective communication via a technical text.

In the following chapter we will look more closely at the field of technical communication and in particular at the nature of user guides. Only by examining and understanding the circumstances under which such documents are produced and the environment in which they are used can be begin to understand what it is a technical translator needs to achieve in order to make translations function as originals.

Chapter 2
Technical Communication

The purpose of this chapter is to provide an overview of technical communication to help translators understand the environment in which technical documents are produced as well as the factors and constraints affecting them. The reason for this is that these same factors and constraints also affect the translation of these texts to a large extent. In this chapter we will also outline the kinds of documents produced and discuss the interrelationships between translation and technical writing. We will then focus on instructional documents in general and software user guides in particular.

Before proceeding it is worth pointing out that in the context of technical documentation, several terms are used to describe what is essentially the same thing. Thus, it is not uncommon to see the terms "user guide", "user's guide", "manual" etc. in use. As Van Laan and Julian (2001:57) point out, the choice of which word to use depends on the particular company, organisation or preferences of the writer and they are largely synonymous. Having said that, they **can** be different if we need them to be different so picking a term and using it consistently is essential. For the purposes of this book, we will refer solely to user guides.

In contrast to other types of writing such as "medical writing", "creative writing" or "legal writing", the field of technical writing, like technical translation, is rather ambiguous in many regards. It is clear that a medical writer writes about medicine and so on but what do technical writers write about? Technical information? This is far too vague to provide any real answers. White (1996:4) defines technical writing as "communicating… specialised information in any field (particularly industry), read by technicians, technical managers, owner-operators of machines, and scientific researchers to perform a certain task". White continues by saying that technical writing is "a means to an end rather than as an end in itself". What makes texts technical is their "utilitarian, specialized focus" (White 1996:12). Technical writing is a form of translation whereby the "abstract-theoretical" is transformed into the "concrete-actual".

However, this definition only goes part of the way towards explaining what technical writing is. Indeed, in the hundreds of years that technical

writing proper has been practised, the entire area has grown so drastically that the term technical writing no longer seems able to contain the wide range of activities it encompasses.

The National Writers' Union (NWU) in the United States maintains that there are at least three different types of writing encompassed by the general term "technical writing" (1998). These areas are as follows:

Technology education Technology education involves writing about technology for non-technical audiences. The products of this type of writing include hardware and software manuals, system administration guides, reports for lay readers, general interest articles, impact statements etc. The NWU points out that writers in this area really only need as much technical knowledge as a typical reader – "namely not a lot".

Traditional technical writing Traditional technical writing involves writing for a technical audience. Writers in this area usually need a strong technical background to produce the repair and maintenance manuals, scientific papers, programming guides and technical specifications required.

Technology marketing Technology marketing is also known as *Marcom*, involves writing sales and promotional materials and corporate communications materials for technology companies. In this area, the NWU states that writers generally only need as much technical background as their audience in order to produce the marketing materials, specifications sheets, brochures, newsletters etc.

The Nature of Technical Documentation

In an attempt to clarify the nature of technical writing and what is actually produced, it would be useful to look at the general characteristics of technical documentation. Markel (2003:7-10) provides a good general overview of technical documentation and its production. To begin with, technical documentation always addresses specific readers. Now we could argue that many non-technical documents are aimed at a particular audience, but technical documents are more specific as regards the audience they are aimed at than most documents. Technical documents are produced taking into account the age profile, job, experience, knowledge, seniority, tasks, problems, aims and objectives. The content, approach, structure, level of detail, style, terminology etc. are all tailored to this profile.

Technical documents help readers solve problems. Markel says that "technical communication is not meant to express a writer's creativity or to

entertain readers; it is intended to help readers learn or do something" (2003:8). As was mentioned earlier, reading technical documentation is generally not an end in itself. People normally read technical documentation because they want to be able to do something else, for example learn how to use software or find out about the design details of a particular device. As Dobrin explains "technical writing adapts technology to the user" (1983:247).

The way in which technical documentation is produced is also important, according to Markel, in defining its nature. The very nature of the company, its culture, goals and organisation, are reflected in the types of documents that company produces. For example, a company with a rigid and formal hierarchy may tend to favour formal and structured memos as a way of communicating between staff rather than informal emails or a simple chat over a cup of coffee.

In producing technical documentation, it is rare for just one person to be responsible for the entire process. Technical communication is a collaborative process involving technical writers, illustrators, editors, subject matter experts (SMEs; pronounced "smee"), designers, illustrators, usability specialists and, of course, translators. This is another important characteristic of technical documentation and it illustrates its complexity. This complexity is also reflected in the tools used in their production. Instead of being produced using a simple word processor, many technical documents are produced using high-end desktop publishing packages and are disseminated via the web in formats such as PDF, HTML and Flash; electronically distributed documentation can even contain animations. All of this makes technical documents more than just printed text on paper. They make the documents more effective and flexible but they require translators to master many of the tools used in to create the documents in the first place.

Another, more immediately obvious characteristic of technical documentation is the way it uses design to improve the appearance and readability of documents. By manipulating the design and layout of a document, we can make it easier for readers to navigate through the document and find the information they need as well as making it easier to understand. As part of this design process, technical documents will often contain a combination of text and graphics to make the document more stimulating and to communicate and reinforce potentially difficult concepts and processes. Graphics in a document make it possible to communicate large amounts of information quickly and clearly and their presence in technical documents is often regarded as essential.

We can categorise technical publications as follows:

- Procedural documents such as assembly instructions, instructions for operation etc.
- Descriptive and explanatory documents such as descriptions of products and services; explanations of processes, concepts etc.; progress reports.
- Persuasive or evaluative documents such as research proposals or engineering projects, product or service evaluations as well as reports recommending actions or policies

- Investigative documents such as reports which are intended to present new knowledge etc.

The first two categories are of most interest to us here because they relate directly to the subject of software documentation of which the user guide is one type. This relationship will be discussed in the following section.

Typical Technical Documents

Technical writers produce a wide range of documentation in an enormous variety of subject areas and industries. From gardening equipment and toys for children to aircraft maintenance manuals, tutorials for word processors as well as ice cream makers and nuclear submarines. The actual types of documents produced can vary according to the subject, the nature of the product and the industry within which the company operates. However, although the focus of this book is on software documentation, or software user guides to be precise, we can identify a number of typical documents that are frequently produced by technical writers and translated by technical translators.

Proposals

Many products and services are come into being as a result of an idea. Ideas need to be proposed to someone in a position to turn this idea into a reality. Proposals are generally an offer to carry out research or to provide a product or service (Markel 2003:483) and may originate outside a company or from within the company. The crucial aspect of proposals is that they are *persuasive* documents. Whether the proposal is for a clinical trial of a new anti-ageing cream or to provide consulting services for air traffic control systems, the ultimate aim is to persuade someone else to agree to an idea. In

order to do this a proposal needs to show that the writer understands the readers' needs, that the proposers can fulfil their promises and that they are committed to doing this (Markel 2003:488). Proposals can be quite challenging for writers and translators because they can frequently involve quite varied and disparate information such as financial information and legal issues in addition to highly technical engineering material. A typical proposal might consist of the following sections:

- Summary

- Introduction
- Proposed Programme
- Qualifications and Experience
- Budget
- Appendices

Reports

At various stages in a product's life, reports will be produced detailing various aspects relating to the product: its development status, viability, commercial success, safety, effectiveness and so on. A report consists of a statement providing facts and information to help readers understand, analyse or take action on some situation or idea (Markel 2003:519). There are three basic types of reports: informational reports, analytical and recommendation reports. Where simple informational reports present information and results, analytical reports provide the same information as well as drawing conclusions based on the facts contained in the report. Recommendation reports build on analytical reports by making various recommendations as to further action etc.

Instructions

Instructions are one of the mainstays of technical communication. Technical writers will probably write more instructional documents than any other type of document (Markel 2003:583). While it is convenient to speak of instructions simply in terms of user guides, there are, in fact, several types of instructional document each of which has its own particular content, format and audience. Repair manuals, for instance, are designed for readers who are not necessarily the actual users of the product. Most likely the readers

will be engineers. As such, a repair manual will not explain how to use the product but rather will provide step-by-step information on diagnosing problems and remedying them. Since the readers can usually be assumed to have quite specialist knowledge, this type of document will have specialised terminology and will assume a high level of expertise. As such, certain information will be implicit, e.g. in the case of a machine, the manual may not state that a certain type of spanner needs to be used to remove an M15 hexagonal bolt.

Similarly, administration guides are also aimed at an advanced audience. Generally written for IT systems, the audience for this type of document are system administrators who are generally IT experts. As such, they are not concerned with the day-to-day tasks of ordinary users but rather in managing, configuring and monitoring the system. An administration guide will provide complex and specialised information relating to the inner workings of the system and will often contain information on diagnosing problems, optimising performance and customising the system.

User guides, however, are aimed at people who need know how to use the product. Often these readers will have no prior knowledge of the product and need to be taught how the product relates to their daily tasks and how to use the product to perform these tasks. Such documents will generally proceed from general to specific information, making sure that all new terms and concepts are adequately explained to ensure that readers do not become confused. The exact content of a user guide may vary depending on the level of users involved, e.g. a beginners guide, an intermediate guide and so on.

One of the primary aims of instructional documents is first and foremost to ensure the safety of the reader and to prevent accidental damage to the product. Instructional documents must anticipate the types of mistakes readers are likely to make and warn them well in advance before they make these mistakes. Once the relevant safety information has been provided, the next task of instructional documents is to ensure that readers know what they need to do and know how to do it. In order to do this, it is essential to understand the audience, what they know and what they need to do.

Software Documentation

Software documentation is a term encompassing a wide range of publications and text types ranging from simple "read-me" files and instructions for use right up to high-level technical specifications. So before we begin

describing these different texts, perhaps we should examine exactly what we understand by *software* and *documentation*.

Software

We can define software as the non-hardware components of a computer. In other words, it can be regarded as the data elements (rather than the equipment) which are fundamental for the operation of computers". Microsoft defines software as "computer programs; instructions that make hardware work" (Microsoft Press 1998). Combining these definitions is perhaps the most enlightening approach in that we can see that software is that part of a computer system which is not physical but which is responsible for making the computer and all of its hardware work. Within the general concept of software, it is possible to distinguish between a number of subtypes:

Systems Software

This type of software includes operating systems which are to a certain extent responsible for "the workings of the computer" (Microsoft Press 1998). This category of software also includes language systems, utility systems and file management systems. They also verify application programs and allow peripheral devices to interface with each other (Houghton-Alico 1985:2).

Applications Software

Essentially, applications perform the tasks for which people use computers. Examples of applications include word processing, spreadsheets, graphics, databases etc.

Hybrid Software

Hybrid software is neither systems software nor application software but rather it contains elements of both types of software. Examples of this type of software include network software, language software which programmers use to write programs, airline reservation systems and online banking systems.

Firmware

Firmware lies somewhere between hardware and software routines stored permanently in read-only memory. This type of software remains unchanged and undamaged even if there is no electrical power supply. Firmware generally consists of basic start-up routines and input/output instructions (Microsoft Press 1998).

Documentation

Documentation can be described as "the set of instructions shipped with a program or a piece of hardware" (Microsoft Press 1998). Houghton-Alico (1985:6) provides a more detailed explanation when she defines documentation as:

> ...the collection of documents that explain, describe, and define the purposes and uses of a particular software program or a system made up of multiple programs.

Documentation generally includes a variety of different documents such as system requirements, set-up instructions, instructions for use, maintenance information etc. All of these types of document can be classified under the following headings:

Technical documentation

As its name suggests, technical documentation provides detailed technical information relating to the initiation, development and operational phases of the lifecycle of a software product. Such information can include functional specifications, design specifications, development procedures as well as test plans and analyses. Technical documentation forms a written technical history of the product from its inception and provides information necessary for the production of user guides, project plans and maintenance (Houghton-Alico 1985:6-8).

Product documentation

This type of documentation provides information relating to product specifications, configuration and set-up details, end-user maintenance and so on (*ibid.*).

User documentation

User documentation is intended to help its readers to use the software in question. In addition, user documentation provides information on other resources for users, explains procedures, defines concepts and generally familiarises the reader with the software and its use. User documentation includes among other things, user guides, online help systems, audio-visual training materials, materials for training courses etc.

The Audience for Software Documentation

The preceding paragraphs provide only a general overview of the types of software documentation. In fact there are numerous types of texts which are produced, each of which has its own particular readership. While we can roughly categorise the readers of technical publications according to the types of documentation listed above, it would be wise to provide an expanded examination of the readership for technical publications.

In general it is usually possible to divide readers of technical publications into "specialists" and "non-specialists". While this may serve as a useful distinction (particularly in light of the document categories above) it is rather too simplified for practical use. The reason for this is that there are many different types of specialist and indeed non-specialist. A specialist may be a programmer who is intimately acquainted with the intricacies and internal workings of the software and how it was programmed. Conversely, a specialist may also be a person who does not know any programming languages but is nonetheless an advanced user of the software. Similarly, one person who is regarded as a non-specialist may have basic knowledge of computers and be fairly comfortable with the general operation of a PC while another non-specialist may have absolutely no understanding of computers and may not have even seen a PC before.

Rather than invent a whole new nomenclature for the various people who read technical publications, it may be useful to look at the practical roles of users and their relationships with the software. White (1996:26) distinguishes between types of reader on such a basis. He proposes the following categories:

Technical Specialist

These include engineers, designers, programmers and testers who need information in order to assemble, develop or operate the software. This type of readership generally has a very high level of technical knowledge and the information they need must be precise, detailed and comprehensive. Documents for this type of readership will contain technical terminology, complex procedures etc.

Production and Promotion Managers

These are the people who decide on the production and marketing strategies for the software. Documents for this readership need to explain the software so that the readership can understand the potential the software can have on a market, what new markets can be targeted, and how production should be carried out.

Administrators

These include supervisors, vice presidents, board members etc. who need to understand the technical aspects of the software in order to draw up benefits, policies, costs and so on.

Clients; Stockholders

These are people who are outside the physical workplace and who are interested in the practical uses and benefits of the software in order to clarify purchasing or investment decisions.

Laypersons

Although we can define this group as the general public who are only interested in general information, it might be more useful to say that laypersons are the general population of users, namely the people for whom the software is a tool in their daily working lives. As such, this group needs clear information on the uses of the software and how to utilise the functions effectively. This information is presented in the form of user guides which contain a broad range of information intended to help readers use the software and find more information if needed. The following section will examine user guides in much greater detail.

Mancuso (1990:58-59) provides a slightly different set of reader categories which is more specific in its definitions. It also presents certain factors and guidelines which Mancuso believes should be taken into account in order to meet the needs of the readers.

Manager: The information should be organised in the form of problem-solution, general-to-specific or most important-to-least important information.

Expert: Documents for this audience should use specialised vocabulary and accepted methodology.

User: Documents for users should present descriptive procedures clearly and they should also use a simple vocabulary.

Technician: Graphics should be used frequently and the guide should not be perfect bound (presumably so that document can be opened flat on a table).

Technical Expert: This readership requires quantitative information to be placed first in the document.

Generic User: Needs abstracts, overviews and summaries to make reading easier. Information should be structured from general-to-specific or most important-to-least important.

Multiple Audiences: Documents produced for multiple audiences should have information layered within sections.

Software User Guides

Mobile phones, video games, digital cameras, MP3 players, word processing software, televisions, x-ray machines, satellite navigation systems, DVD players and industrial process control systems. A reliance on semiconductors notwithstanding, a common theme linking this diverse range of products is that they are all accompanied by some sort of user guide. More specifically, they invariably include a software user guide.

A common misconception about software user guides is that they are written only by software companies for software products. In reality, however, any company that produces software - even if it is only as a supplement to the company's main product - will produce software user guides (Van Laan & Julian 2001:4).

Despite the wide-scale production of user guides, especially in the software and technology industries, it appears quantity has not translated into quality. Indeed, poor or inadequate user guides are a widely acknowledged problem in industry (Brockman 1990:1). That there exists such inadequate documentation is disturbing, especially when we consider that aside from certain legal implications, the quality of user guides can spell success or failure for a product or even for a company. One such documented example refers to the catastrophic losses incurred in 1983 by a company called Coleco. This company lost a staggering US$45 million in the final three months of 1983 as thousands of irate customers returned the Coleco Adam computer, citing the terrible user guide as the problem (Brockman 1990:13). Stories like this are numerous and it is clear that user guides can help improve sales and create loyal customers (as was the case with Apple computers in the 1970s and 1980s).

But high quality user guides are not just useful ways of gaining new customers. As products become more and more sophisticated and complex, it is essential that quality user documentation is available to help users exploit the full range of functions offered by the product. Companies frequently spend vast sums of money on products of which only a fraction of the functions will ever be used.

It is possible that the problems of poor documentation are due to the simple fact that companies do not always understand the purpose of user

guides and those who produce them, not to mention the needs of the customers who use the user guides.

User guides are, in effect, an interface between computer systems and their human users. In producing user guides, technical communicators need to act as an intermediary between the software and the users. To be successful in producing user guides, it is essential not only to understand the product in detail, but also to understand the humans who will use it. Coe (1996:2) states that technical communicators design information for users and there is “a covenant of trust” between communicator and user. This, she maintains, is the basis for human factors in user documentation.

Admittedly, in recent years, the quality of user guides has improved steadily. Yet there are huge numbers of people who simply do not read user guides no matter how complex the products they want to use. In fact, it seems that sometimes the more complex the product, the less likely some people will be to read the user guide. While we can attribute this to an expectation by users that modern products are generally intuitive and self-explanatory, it is more likely that we are dealing with people who Coe claims have “lost their trust of technical communications” (1996:3). For them, “past experiences may have destroyed their trust and colored their approach to and use of” technical information presented in, for example, user guides (*ibid.*).

Here, the problem facing user guides is more serious and more difficult than simply teaching users how to use a product. Rather, the task is to re-establish contact and trust with users and persuade them to read and trust user guides. These users frequently have just reason to be wary of user guides because previous experiences have left them feeling frustrated, confused or just plain stupid. An interesting discussion of this issue is provided by Schriver (1997:214-222) who cites feelings of confusion and incompetence among users as a result of inadequate instructions. Interestingly, users generally blame themselves for their inability to follow instructions for products (*ibid*:222).

What is a User Guide?

Major defines a user guide as…

> a summary of all related resources […] written for everyone who might use these resources. It serves as a system of pointers to more detailed information, such as references manuals, tutorials, and standard operating procedures. (Major 1985:122)

At first glance, one would be forgiven for thinking that a user guide was little more than a directory or compendium of existing documentation or some form of super product encyclopaedia. However, Major clarifies this when discussing documentation policy within a particular company. With an abundance of tutorials, how-to documents etc. a user guide was created in order to organise and arrange the existing documents. So rather than being a mere directory, Major explains that the user guide actually consists of many of these resources while providing references to others. He goes on to say that a user guide is a resource which not only points to all available documentation but also guides the reader through a general background and lays the foundations upon which "assumptions can be made about the level of comprehension readers have when they reach a given point in the total documentation picture" (*ibid.*).

Weiss (1985:4) provides a simpler definition of a user guide (or manual to use Weiss' terminology): "a user manual is – or should be – a tool that helps its readers get full benefit from the system." The guide, he maintains, is intended to compensate for the fact that software and information technology are often difficult and unfriendly to the user. This is also true by virtue of the fact that there is a limit to what can be learned autonomously and intuitively without assistance. Weiss continues to say

> …users may become interested in the inner workings of computer technology. But this does not alter the basic idea: Users are people who want something bigger than, and outside of the particular device (*ibid.*)

It is clear from the preceding definition that users use software as a means to an end – a tool to help them do something else.

The Function of User Guides

Given that users want to learn how to use the software as quickly as possible in order to carry out their main day-to-day work, a logical starting point when discussing user needs is the expectation that the user guide will provide clear and unambiguous instructions. This makes the guide easy to understand and ensures that the information is correctly interpreted by the reader. A key factor in achieving this clarity is the definition of technical terms and new concepts as and when they are used.

Users also need guides to be structured. They need to know what the guide covers and what it does not. Perhaps more importantly, the information must be organised according to what users want to do. Price puts it as follows:

> Organize by what I want to do – the tasks I want to accomplish – not what some programmer or engineer thinks. [...] don't make me jump all around the manual just to find one fact (Price 1984:8)

Users need to be taught – not just presented with information. Readers need to be provided with access not just to the documentation but also to the specific information contained in it. They need to be guided through the basic functions of the software on a step-by-step basis and the guide should tell users what to do and what they should see at any given moment. The whole point of a user guide is to convey enough information to users to allow them to perform tasks as quickly and as easily as possible and with a minimum of confusion and effort.

The ability of users to quickly learn how to use the software is also affected by the usability of the guide. Users need to find information quickly and easily. Summaries, glossaries, indices and clear, informative tables of contents speed up access to and retrieval of the necessary information. Appropriate language usage, which also reinforces the clarity of the guide, means users will not have to waste valuable time deciphering ambiguous sentences or phrases.

Above all, users need the information in user guides to be correct. The information must accurately reflect facts – any errors will damage the users' confidence both in themselves and in the guide and will ultimately impair if not ruin the learning process.

The previous paragraphs have hinted at the functions of user guides but they provided only a very cursory insight. According to Weiss (1985:16),

the primary goal of user guides is to control the reader and the communicative action. Weiss maintains that

> …to communicate well we must respect the independence and intelligence of the readers, but must not rely on them. […] For user and operations manuals, the best strategy for writers is to adapt to the weaknesses in typical readers, to assume control of the communication. (*ibid.*)

Weiss is first to admit that any reference to controlling the audience or communication can raise strong ethical criticisms. However, he justifies this by saying that while we do not fully understand how people read or comprehend, we do have some knowledge about what distracts them or causes interference in the reading and comprehension processes. Thus, in removing sources of "noise and error" (*ibid.*) and things which we know will interfere with the correct and effective use of the guide we are in a sense assuming control over the reader and the communicative act.

If, for example, to quote Weiss a guide is little more than a collection of facts and pieces of knowledge, the effectiveness of the guide depends on how well the reader processes, sorts and assimilates the information. If, on the other hand, the guide is "engineered to suit the interests and abilities of the reader, then the user is to some extent prevented from misusing the material" (*ibid.*). In this regard it would, perhaps, be better to rephrase the goal of user guides and say that they should *guide* the reader and the communicative act by limiting what the reader can do with the user guide and limit the use of the guide to the intended one.

The functions of the information contained in a user guide, presented taking the preceding paragraphs into account, can be summarised as follows:

Help the user get started

The guide should introduce the functions and benefits of the software, demonstrate how to install and configure the software and teach users how to use elementary functions. It should also provide the users with precautionary information and warnings about potential errors (Weiss 1985:5).

Help increase productivity

Having shown users how to get started, the user guide should then take them through the advanced functions of the software, teach users to utilise

the software more effectively, introduce shortcuts and how to apply the software to the users' own particular needs (*ibid.*).

Troubleshooting

Once users are at an advanced stage in the use of the software, they need to know what to do when things go wrong. In order to do this they need to be able to identify problems and find appropriate solutions, distinguish between problems which can be rectified by the users themselves and those problems which require expert assistance (*ibid.*).

International Readers

With the increasing level of globalisation of world markets and the ever shrinking nature of the world as a result of modern communications, transport and multinational companies, there have been quite significant changes in the way international markets and communications work. An increasing number of companies are using English as a working language (see Byrne 1999:37). Indeed, given what White (1996:32) calls the "globally interactive, multicultural and multinational" nature of the modern workplace it is clear that there is a perceived need for some form of *lingua franca*.

It has been estimated that there are some 1,200 million speakers of English in the world but only 350 million of them are native speakers (Byrne *ibid.*). This means that there is a huge English-speaking readership for technical documents and it is important that we take into account the cultural and linguistic differences associated with these readers. While Byrne (*ibid.*) concentrates solely on the need for linguistic compatibility of documents for different English-speaking audiences, White (*ibid.*) emphasises certain cultural differences. He maintains that the way in which information is conveyed can vary quite significantly from culture to culture.

> What might be complimentary to one audience may be offensive to another. What might be crystal clear to one culture may be incomprehensible to another.

The way in which information is processed is often different for different cultures. This can affect the way a document is structured or even the way information is packaged and structured on a paragraph or sentence level (cf. Gerzymisch-Arbogast 1993).

On a personal level, certain forms of address used in one culture may, according to White, be inappropriate and considered rude in other cultures.

Rather than the mere use of standard formalities, the cultural requirements may mean that the way instructions are given may need to be modified. Of course it would be virtually impossible to avoid every potential cultural and linguistic pitfall in a user guide. However, as part of the audience definition stage of the production or translation of a user guide this information can prove valuable and worthwhile for the specific audiences being targeted.

Other Functions

Although the function of user guides is to educate and guide the readers, Weiss (1985:12) divides this *instruction* function into a further three sub-functions which may be attributed to individual user guides to varying extents. This additional sub-categorisation is by no means exhaustive and it is conceivable that several more functions could be attributed to user guides. The following paragraphs are intended merely for illustrative purposes.

Tutorial

According to Weiss (*ibid.*), tutorials are "instructional materials intended to train neophyte users." Typically, such materials follow a simple task-orientated format aimed at helping absolute beginners take the first crucial steps in using a product. The subjects covered in tutorials tend to be basic, bottom-up concepts. Tutorials thus can be described as providing "learning support" as opposed to the purely procedural support provided by instructions.

Demonstration

Weiss defines demonstrations as "materials aimed at teaching a process or activity to a competent or experienced reader." Such materials literally teach users by showing them how to do something. Demonstrations show entire processes in a top-down manner because the users already have an understanding of the basics, what they need to know is how to put these basic steps together to complete a significant task.

Reference

This type of information is aimed at advanced users who, according to Weiss "know what they need to know." Rather than explaining why certain information is important and how it relates to other information, the

reference material is simply a compressed version of key facts relating to the product.

Motivation

Often the people who read user guides are under no obligation to actually follow the instructions provided in a user guide. It is also the case that they may not even want to read the user guide in the first place. Schriver maintains:

> Bad experiences with documents influence not only our thinking about documents we've used but also our attitudes about documents we have not yet seen. People appear to generalize their bad experiences (Schriver 1997:4).

In such instances perhaps one of the most important functions of user guides (apart from pure instruction) is to motivate the reader, to make it seem worthwhile for the reader to read and follow the guide. Examples of such "motivating" guides would include the "*…For Dummies*" books published by *IDG books*, which take a range of different subjects and present them in an easy to follow way for people who, for whatever reason (often because the subject is perceived to be too difficult), may be reluctant to learn.

With this in mind it is, according to Borowick (1996:176), important for the writer to explain the importance of the guide and the instructions it contains. Borowick maintains that a conversational tone, among other things, will encourage the readers to co-operate. Borowick begins by saying that it is necessary for the writer to discuss at the very outset the "desire to achieve a common goal" (*ibid.*) or to express concern for the satisfaction and success of the reader as a user. The results and benefits should, therefore, be explained to the readers. The explanations or justifications for certain instructions must be clearly separated from the actual instructions. Borowick says this can be achieved using italics, parentheses or other "literary mechanical devices."

Software User Guides & Quality

Previous sections have alluded to various factors which can influence the quality of a user guide but this only part of the picture. The reality of user

guides means that there are numerous ways in which the quality of user guides can be affected and these cover several areas including linguistic issues, organisational factors, social considerations involving users, technical and visual features to name, to provide a few examples. The best way of understanding user guides and how to improve them is to look at the various factors that feature in their production. In the following section we will examine the problem of poor user guides and explore a wide range of issues affecting their production. This will help us to understand user guides and how we can improve their quality.

The Cost of Inadequate User Guides

The consequences of inadequate user guides should not be underestimated. Approximately half of all product returns and complaints in Germany arise as a result of bad instructions. Frequently, customers end up damaging products themselves because they lose patience with bad user guides and resort to "creative" and sometimes unorthodox methods to get their products to work (Cognitas 2003a). The resulting damage and compensation amounts to some €500 million each year in Germany alone (*ibid.*). This is due in large part to changes in European Union laws governing product liability. European Resolution *C411* states that

> …inadequate operating instructions may affect the presentation of products and may be a factor to be taken into account together with all other pertinent circumstances in considering whether goods are defective (Council of the European Union 1998:1)

This resolution goes on to say that in the light of the wider range of products and the advances being made in technology…

> …operating instructions for technical consumer goods are often perceived by consumers as inadequate, both because they are unclear and present language difficulties, owing to faulty translations or to the use of terms which are too complex, and because they lack structure and have inadequate content.

Tackling the Problem of Poor User Guides

The requirements in this resolution have filtered down into national laws. For example, in early 2002, Germany's product liability law (*Produkthaftungsgesetz*) was overhauled with the result that user guides are regarded as a part of the product and as such, any defects or faults they contain are regarded as product defects which can result in the rejection or withdrawal of the product (Heino 1992:111). In addition, where a product is damaged or destroyed as a result of a user following instructions contained in an inadequate user guide, the manufacturer or retailer is obliged to provide a replacement (Cognitas 2003b).

To counteract the problem of poor user guides, the European Union has codified what it believes to be the essential characteristics of "good" user guides. Resolution C411 (Council of the European Union 1998) sets out, among other things, a series of criteria under the following headings, which the Council of Europe believes will make for more effective user documentation:

- *Development of documentation*: all relevant laws, standards and guidelines should be consulted and the document must comply with their requirements
- *Content of documents*: the content of documents should be structured logically and reflect real use; warnings and cautionary information must be clearly distinguishable from the main document content
- *Separate documents for different models*: unless the procedures involved in using functions are identical, separate documents must be produced for different models or variations of products
- *Safety and warning instructions*: must be clear and easily accessible
- *Document language*: user documentation must be available in a user's own language
- *Style and layout*: should ensure clear and readable documents

The overall aim of this is to produce high quality documentation which will provide customers with "adequate user information to ensure proper and complete use of the product" (Council of the European Union 1998:1). Other regulatory and legislative tools governing the production and provision of user guides include:

- *EN 62079*. "Preparation of instructions - Structuring, content and presentation"

- *EN 292-2* “Safety of machinery. Basic concepts, general principles for design”
- *VDI 4500-2* “Technical documentation - Internal technical product documentation”

These standards and guidelines have gone some way towards ensuring better user guides are provided to users. One initiative, based in part on these regulations, is the *DOCcert* certification scheme developed in 1993 by *tekom* and *TÜV* in Germany (Jung and Becker 2003). This is a quality assurance and certification programme aimed at ensuring documentation is effective, complete and facilitates the safe use of products. The certification process tests documentation for comprehensibility, completeness and safety and takes place in three stages.

The first stage involves examining relevant laws, standards and guidelines such as those mentioned above and ensuring that the documentation complies with their requirements. The second stage involves testing the documentation on the basis of a series of criteria such as accuracy, comprehensibility, layout, readability etc. The final stage involves hands-on usability testing with users. Successful documentation is then certified by *TÜV* and can bear the *TÜV*-approved logo.

These initiatives notwithstanding, it is clear that work on improving the quality of user guides is far from complete and that there are still countless inadequate user guides in circulation. A study conducted by the German computer magazine *ComputerBild* in 1999 examined 60 user guides from a range of well-known companies and found that 35 could be regarded as “faulty” and could result in complaints or claims for compensation (ComputerBild 1999:16). Using the *DOCcert* test procedures and criteria, the investigators found that only 4 of the user guides passed the stringent requirements.

The obvious need to overhaul the way in which user guides are produced has serious implications for vast numbers of technical communicators across the world. Up until now, we have referred to technical communicators as being responsible for the production of user guides. While traditionally *technical communication* would be taken to mean technical writers alone, the industry and nature of the work have developed to a point where technical communication includes the work of technical writers, illustrators, technical translators, editors and web designers (Van Laan & Julian 2001:5). Indeed, many professional technical communication associations explicitly include these roles under the umbrella term of *technical communication*. Given the fact that according to Council of the European Union *Resolution*

C411 "customers are entitled to manuals produced in their own language" (Council of the European Union 1998:3), it is clear that "translation work [is] an integral part of the process of creating technical documentation" (Budin 2000).

Technical writing and technical translation are inextricably linked with regard to user guides. As such, any discussion of user guide quality must take this relationship into account, not least because translation is explicitly identified in the aforementioned European Directive as a potential cause of problems in user guides (Council of the European Union 1998: 1).

However, the problems of poor documentation are sometimes best tackled at ground level by the people who actually write and translate them. This is the reason for this book: to help understand the problem and issues which affect it as well as how translators can contribute to the production of high quality technical documentation.

What Makes A Good User Guide?

Unfortunately, many of the factors affecting the quality of user guides are present even before the technical writer has put pen to paper so to speak and long before the text reaches the translator. These problems are frequently impossible to correct once production of the user guide has started so it is essential that they be addressed from the outset.

One of the most significant problems facing user guides is the fact that they are frequently a casualty of poor planning and excessively tight schedules. When developing software, the primary focus for many companies is to make sure the product has been built and that it is working. In the grand scheme of things, user guides usually figure way down the list of priorities. Translating the user guides is an even less pressing priority. Admittedly, the issue of excessively tight deadlines is intrinsic to the development process and cannot easily be remedied because documentation cannot be completed until the software has been finalised and translation cannot begin until the documentation is complete. However, it is a truism in the translation industry that translating documentation is frequently left until the very last minute. This places unnecessary pressure on technical writers and technical translators and eats into the amount of time available for research, testing and refining texts.

Setting aside organisational problems such as scheduling, there are a range of areas which directly affect the quality and effectiveness of user guides. Traditionally, issues such as audience definition, document style and

design were ignored and user guides were often produced as an after-thought. A turning point in the development of user guides came in 1978 when Apple Computer Inc. launched a new breed of user guide to accompany its new computer. The following is an extract from a press release issued by Apple Computer Inc. to mark the launch of the guide:

> Created with foremost concern for the reader, the book assumes no prior background in programming or computers. Programming is explained in everyday English with no computer jargon used. Moreover, with scrupulous attention to detail, the book introduces the whole computer to the reader. Thus unlike programming manuals that solely teach a language, this book teaches a language in the context of the computer in which it will be executed.
> Another contrast with stereotyped programming manuals is the book's graphic illustration and **literary** style. Using a two-color process with text in black, significant information is highlighted in eye-catching green. Moreover, to illustrate displays, actual television displays are used to ensure the reader that observations on the television monitor will be the same as those within the book. Furthermore, the manual's informal, slightly humorous style, makes the book truly enjoyable to read. [emphasis my own] (Apple Inc. 1978)

One important point which should be made here is that the claim regarding the book's **literary** style should not be taken at face value. In this case, it does not mean literary in the traditional sense. Instead, this book used language devices and styles in stark contrast to previous documents with their dry, machine-like and "robotic" language.

This development set a new standard in the production of user guides and manufacturers soon realised that good user guides could actually win them customers (Price 1984:7). And so, terms like usability and design became part of the day-to-day vocabulary of technical writers and documentation departments as user guides came to be treated like devices (Weiss 1985:10-11). Companies learned that documentation needed to be well written, easy to understand, well laid out and presented, enjoyable and colourful. Of course there are countless factors which play a part in the quality of user guides and it would be impractical to discuss each and every one of them here. However, it is possible to deal with a representative selection of factors grouped according to the general areas of user guides they relate to, namely: *Appearance*, *Content*, *Structure* and *Language*.

Appearance

The appearance of user guides is an essential part of the communicative effect of user guides and it conveys information to users before they have read a single word. Appearance includes anything from the arrangement of text on a page to the way a guide is bound – all of which affect the effectiveness and usability of a user guide. Houghton-Alico (1985:59) maintains that in order to produce quality software user guides we need to devote as much time and effort to format and design as we do to content and writing style.

Apart from making a user guide more aesthetically pleasing, the way in which the pages of a user guide are designed and laid out plays a crucial role in how readers find and assimilate information and can even determine the environments in which the guide is used. According to Houghton-Alico (1985:59), each page should invite the reader to read the page, to become involved in the user guide. It is also essential that the appearance of a user guide is consistent with the look and feel of other documents produced by the company as it creates a more professional impression and serves to improve confidence in the product; conversely inconsistently presented documents impair the image of the company and can also shake a reader's confidence in the quality of the product and the information contained in the document.

The design of each page should, according to D'Agenais & Carruthers (1985:48-50) take the following criteria into account:

Simplicity

The design of the page should not be distracting or visually "busy." The information should be immediately apparent to the reader.

Retrievability

The page should have enough information on it to facilitate the immediate identification of the subject matter. For example, the document title or chapter name should be printed in the header or footer.

Flexibility

The design must be able to accommodate all variable data such as department names, people, etc. without any formatting changes.

Readability

Schriver (1997:263) maintains that the best length for a line of printed text is approximately 40-70 characters or 8-12 words as it is easier to read. A similar point is made by D'Agenais & Carruthers (1985:101) who say that a

sentence should be between 10-15 words in length. The page should also have plenty of white space (see below).

Functionality

The page margins should allow for binding and double-sided printing and the headers/footers should use as little space as possible in order to maximise the amount of space available for the actual text of the user guide.

White Space

Perhaps one of the most important factors in page design is the relationship between printed matter and white space. Borowick (1996:132) defines white space as "any part of a page that is blank and used to separate ideas." This relationship is known as spacing and includes spaces between lines, paragraphs and margins.

Margins

Margins should be wide enough not just to facilitate binding (Austin & Dodd 1985:50) but to increase the amount of white space and "prevent the reader's eyes from running off the end of the page" (Borowick 1996:130). It is generally agreed that the page margins should be at least 1 inch on all sides with an additional 0.5 inch for the inside margin (Borowick 1996:130) although D'Agenais & Carruthers (1985:185) suggest a 0.7 inch margin.

Columns

White space can be increased by using a two column format where the left column is used for headings and the right is used for body text (Mancuso 1990:139). The following diagram illustrates this concept.

Figure 1: Two Column Page Layout

Paragraph Spacing

Paragraph spacing refers to the way separate paragraphs are presented as separate "chunks" of information using white space. According to Mancuso (1990:133) paragraphing and paragraph spacing reduce the amount of fatigue experienced by readers. As regards the actual amount of space between paragraphs, Mancuso (1990:139) recommends skipping lines. Houghton-Alico is more specific in this regard and says that the spacing between paragraphs should be 50% more than the space between lines.

Line Spacing

The spacing between lines of text is generally 2 points (see *Typography and Formatting* on page 73 below) larger than the size of the fonts used in the text. So, for example, if the font size is 10-point, the space between the lines will be 12-point. Mancuso (1990:139) recommends using one and a half or double line spacing between all lines of text. This is echoed by Borowick (1996:131) who states that single line spacing looks crowded and is difficult to read. One and a half spacing, according to Borowick (*ibid.*), allows the reader's eyes to "drop naturally at the end of each line to the beginning of the next line." Double spacing, he continues, is excessive and makes for uncomfortable reading.

Paper Size

One rather obvious factor affecting the amount of white space is the actual size of the paper being used. Quite simply, a larger page allows for more white space once the page margins have been incorporated and the pages have been bound. While standard paper sizes (e.g. DIN A4 or US Letter) are preferable, it is possible to use different sizes. However, using non-standard paper sizes can significantly increase production costs and may also affect how easy a user guide is to use (e.g. it is difficult to use a large and bulky guide while seated at a computer workstation or while working in a factory production hall).

Typography and Formatting

The fonts and formatting used on text in a user guide play an essential role in the effectiveness and usability of a user guide. The main requirement when choosing fonts is that they are clear and consistent (Austin & Dodd 1985:50; White 1996:204). Fonts can be defined using three basic characteristics (White 1996:203):

Font Type

Font type refers to the general design of the font. Fonts are divided into what are known as *serif* and *sans-serif* fonts. Serifs are small strokes or continuations at the top and bottom of individual characters or letters. Serif fonts, such as *Times New Roman* and *Bembo*, feature these small strokes; sans-serif fonts, such as *Arial* and *Century Gothic*, do not (see Figure 2).

Whereas sans-serif fonts are clearer, more visually striking and generally ideal for titles and headings, serif fonts are easier to read over a prolonged period of time and are suited for use on body text. This is because the serifs allow the reader's eye to follow the reading line better by leading from one letter to the next (D'Agenais & Carruthers 1985:185).

Font Styles

Font styles refer to variations or alternative forms of a font style such as *italics* or **bold**. Italics can be used to highlight definitions, to emphasize words, to denote trademarks or foreign words. Italics can also be used to highlight a word the first time it is used. Bold is used to emphasize words or to draw attention to certain information.

Font Size

Font size is measured in points. Sizes generally range from 5-point for small text right up to 72 for headlines. Standard text normally uses 10 or 12 point fonts.

Figure 2: Serif and Sans-Serif Fonts

Text Alignment

The amount of white space and thus the readability of the text are affected by the way text is aligned on the page. There are four principle types of text alignment. *Centre alignment* is useful for headings and titles. Right alignment involves placing the end of each line of text flush along the right-hand margin. Conversely, left alignment involves the start of each line of text being placed flush along the left-hand margin. The result is what is known as "ragged right" edge along the right-hand side of the text caused by the uneven line length (Mancuso 1990:144; D'Agenais & Carruthers 1985:185).

Justified alignment involves "stretching" each line so that the start of each line is placed flush along the left-hand margin and the end of each line is placed flush along the right-hand margin. This type of alignment is achieved by adding additional spaces between certain words in order to pad the length of shorter lines.

This decision to use one type of alignment over another, in particular ragged-right versus justified is not without its problems. Where some studies, such as Gregory & Poulton (1970), have shown that may be more difficult to read, others such as Fabrizio *et al.* (1967) and Hartley & Burnhill (1971) found no difference in reading speed or comprehension between ragged-right and justified text.

While justified text looks neater on a page, it is more difficult to read than ragged-right text if there are "rivers" of white space running down the page. The problem here is that the additional and often unpredictable spaces added to the lines make it difficult for the eye to proceed at a constant

rate along the line of text. Each time an extra long space is encountered, the eye pauses briefly. Such pauses or hesitations are known as "fixations." In contrast, the even spacing between words in left aligned text eliminates this and the uneven line length adds variety for the reader and makes the text less monotonous. It is widely recommended that body text be left aligned (D'Agenais & Carruthers 1985:185).

Delivery

Delivery is used to refer to the physical means of producing and distributing user guides. The way the user guide is delivered affects the way the guide is used, how frequently it is used as well as how easy it is to use.

In recent years, the lines between documentation and software have become somewhat blurred with the advent of single-source, multi-channel publishing whereby a single stock of text is produced for use in a variety of media such as printed documentation or online help etc. In addition, there has been a trend which has seen manufacturers provide documentation in electronic form only, for example, in the form of PDF files. Some software does not come with what is traditionally regarded as documentation, favouring instead, comprehensive help systems with complex interfaces which allow task-specific items of information to be accessed directly from within the software.

This has been justified for a number of reasons, most notably those relating to cost. There are many arguments that producing print documentation is considerably more expensive than producing online documentation. However, on closer examination, such arguments are less than convincing. Starr (2001) and Curwen (2002) maintain that the question of costs is less to do with actual costs and more to do with ***who*** bears the costs of printed documentation. They argue that while manufacturers generally escape relatively unscathed by distributing documentation on a "CD-ROM that you know cost them around one dollar to manufacture" (Starr 2001), users ultimately end up printing the documentation on an inkjet or laser printer; the cost to the user can frequently be double the cost the manufacturer would have incurred.

But the real reason why justifications for the proliferation of online documentation are inadequate lies in the reason why users feel the need to print off sometimes entire chapters of the online documentation: they frequently find it too hard to read from a computer screen. The fact that it is commonly believed that users take 20-30% longer to read from a screen than from a page (Curwen 2002) is based on established research by the likes of Dillon (1992). So while the provision of printed documentation is

ostensibly assuming less importance for software manufacturers, for users, they continue to be essential.

However, the range of visual devices is not restricted to screenshots. Other devices include tables, graphics, graphs, charts and diagrams.

Graphics

Graphics are a fundamental component of technical documents in general

and user guides in particular. The term "graphics" means graphs, charts, pictures, photographs, icons, diagrams and drawings. The target audience is a key factor in deciding what type of graphic to use. Graphics are an aid to communicating ideas and concepts clearly and quickly rather than using paragraphs of text explaining the same concept. When it comes to actually implementing graphics in a user guide, there are a number of generic practical guidelines which should be considered:

- Graphics must be referenced sequentially in a document (Borowick 1996:102)
- A graphic should appear on the same page as its reference in the body of the text. A graphic should, according to Borowick (*ibid.*) never come before its first reference in the text. Sometimes, however, documents may have all graphics collected at the end of the document. Nevertheless, they should be numbered, captioned and referenced.

Screenshots

In addition to textual information, pictures of the software interface can also be used in user guides. These pictures are known as *screenshots* or *screen grabs* and an example is given in Figure 3 below.

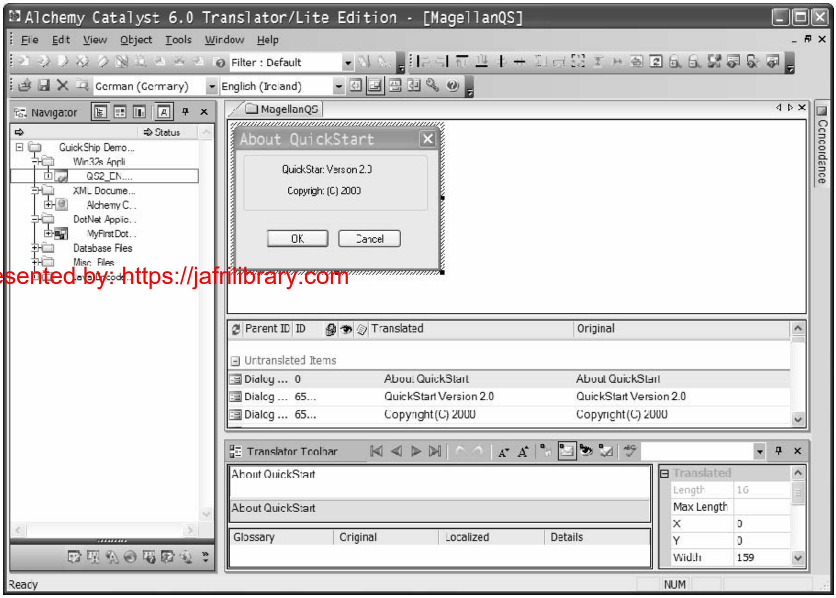

Figure 3: Screenshot from Alchemy Catalyst 6.0 Translator Edition (Courtesy of Alchemy Software)

Screenshots present a unique set of problems for translators in that they are often taken from software which has not been localized yet, which has been partly localized or which is still undergoing development and is liable to change. Consequently, care must be taken to ensure that any references to items in a screenshot are accurately translated and that they are still present in the final version of the software. Esselink provides a detailed and enlightening discussion of the challenges involved with screenshots (2000:349-355).

Other types graphical devices used in user guides include graphs which are used to depict trends or relationships between variables. They are a visual representation of information and are a good way of conveying information quickly. Drawings and diagrams are visual reconstructions of an object and may be representational or abstract. They help the reader to visualise physical objects.

Photographs are the most realistic way of representing an object, but their use is not always advisable because, according to White (1996:236) it is "difficult to be selective about the visual information you wish to present". For instance, if a user's attention is to be drawn to a particular part of a computer's hardware (e.g the motherboard), a photo of the entire internal

mechanism of the computer will leave the reader searching for the part in question. A drawing or diagram of the particular area would be more effective because, depending on how the diagram is drawn, it is possible to focus attention onto the specific part while eliminating distracting visual information.

Photos also raise the issue of print quality and printing costs as they must be of a sufficiently high quality in order to ensure that they are clear and effective. However, the use of high-quality photos also raises the cost of printing the documents.

Structure

The way in which a user guide is structured is fundamental to the success of the user guide and the ease with which readers can use it. Far from being a vast, disordered repository where information is dumped and through which users have to sift in order to find the information they need, a good user guide presents users with the information they need and, just as importantly, when they need it. The way in which information is structured in a user guide can depend on such things as the nature of the product being documented, the background of the audience, the tasks the audience needs to perform and so on. If a user is bombarded with information, the likelihood that they will find the information they need, much less understand it, is greatly reduced. To combat this, we can structure user guides in a number of ways:

- *Chronological*: this structure is used, for example, to describe steps or tasks that need to be carried out in sequence
- *General-to-specific*: this can be used to describe, for example, the background, preparations and safety precautions needed before providing step-by-step instructions
- *Problem-Methods-Solutions*: this type of structure is particularly useful for providing tutorials, maintenance information and trouble-shooting sections or guides and it presents information according to specific problems and the measures needed to resolve them
- *Cause-Effect*: this structure can be used in conjunction with the *problem-methods-solutions* approach for trouble-shooting sections or it can be used to describe the components of a product, e.g. the interface buttons and toolbars.

According to Weiss (1985:50), user guides are structured in that they represent a top-down approach to a particular task, e.g. providing information on a software application. By this we mean the user guide starts with the "big picture", the largest possible overview and then progressively and systematically adds more and more detailed information. While it is generally held that this top-down approach involves breaking concepts into smaller and smaller ideas (cf. D'Agenais & Carruthers 1985:68-9) – a process known as "decomposition" – it is, in fact, only a small part of the structured approach to user guide design.

First and foremost, a user guide needs to provide a broad overview in order to explain how the various constituent ideas and tasks relate to each other. This can be explained as follows: using the example of a piece of website management software, the big picture is creating a website but in order to do this we need to create the web pages. We then need to create directories in which to store the pages and, once we have done this, we then need to know how to upload the pages and directories to the web server.

As a result, we can see that we need to create the directories in order to store our web pages effectively and to allow navigation but these pages and directories cannot become a website unless we know how to upload them. Similarly, we cannot upload a site if it is not organised into directories. Neither is there any point in uploading directories if they do not contain web pages. So in addition to breaking ideas down into smaller tasks, we need to tell readers how these tasks all relate to each other. Thus, a user guide will generally consist of many small units or modules, all of which are connected in a way that will make it clear to readers what they are learning and why.

Modules

The majority of technical documents and user guides in particular consist of small, independent sections (Markel 2003:15). Weiss (1985:52) describes these sections as modules which are small, independent, functional entities, which are components of some larger entity. Modules are functional in that they do something, they perform some task. What is more, they perform a complete task with a clear beginning and end.

Modules Are Independent

Modules do not depend on their context. Since there is a clear beginning and end, the module can function in isolation and may even perform that function in more than one situation (e.g. a module explaining how to save

a file under a different name in Microsoft Excel would be equally effective at explaining how to copy a document to another location using *Microsoft Word*). Thus, modules can, according to Weiss (*ibid.*), become part of a "library of reusable modules". And so these modules eventually resemble a set of stock modules which can be picked and mixed by writers.

We know that modules perform only one function (even though this function can be performed in a number of contexts). But this does not give us a clear understanding of how small a module should be. It could be argued that it is impossible to say with any degree of certainty how small is small. As modules get larger, they begin to incorporate more functions and begin to lose cohesiveness. On the other hand, however, as they get smaller, the links between modules become more complicated. So in the absence of any real alternative, we find ourselves faced with a balancing act of "not too much, but just enough". A practical approach would be to concentrate on one "task" as opposed to a function.

Navigating Modules

A table of contents (TOC) is a map of the document which allows readers to find their way through the modules in the document. A TOC should at the very least include chapter titles and the first level of headings from each chapter as well as the page numbers for each title and heading. However, depending on the exact content of each chapter, additional levels of headings may be necessary.

There are a number of other strategies which are of use not only to writers but also to translators. The first of these is that chapter titles and level 1 headings should make sense to beginners. They are generally quite broad, for example, "*Installing…*", "*Configuring…*", "*Using the Software*", "*Solving Problems*" etc. However, the headings within chapters become more specific and tell readers what they can do or what the section can do for them. One way of reinforcing the notion that a reader can do or achieve something is to use verbs in headings as opposed to "a bunch of nouns, which look like a list of topics to study" (Price 1984:65). Price also suggests phrasing headings as questions as a way of making them more interesting and informative. It could be argued that this is because the questions may actually reflect questions the reader may be asking. An interesting examination of headings is provided by Gerzymisch-Arbogast (1993) who argues that texts are composed of information that is *given* or *new* relative to the reader. The relative amounts of this information and the way it is sequenced are, she says, dependent on the author-reader contract governing a particular document and must be changed to suit the needs of a new

audience. Thus, in an English document, headings generally contain either given information or a combination of given/new information.

In addition to making headings informative and clear, the very nature of the TOC which is generated from the actual headings used in the text places constraints on the length of titles. Short titles, in addition to looking neater on a page, are actually easier to read (Price 1984:67). A reader should not have to scan a heading a number of times in order to elicit the information. Price (*ibid.*) recommends that titles should not take up a full line.

Overviews and Summaries

As readers progress through a user guide or even when they dip into a user guide to read a chapter, it is important to tell them what they will find in a particular chapter or large section. This helps them to decide whether the particular section is what they are looking for. Overviews can also help readers absorb, understand, learn, and remember information more easily (Foss *et al.* 1981).

Price (1984:72) and D'Agenais and Carruthers (1985:90) state that every chapter and large section should have an overview. Consistent with the motivational function described on page 64, overviews explain to readers why they should read a particular chapter or section, what they will be able to do afterwards and what they will be able to achieve with this knowledge.

In general, a typical overview will explain what readers have learned so far (providing they have been reading the user guide in sequence) and how this particular chapter builds on previous chapters. The overview tells the reader what will and will not be covered in the chapter and provides a broad idea of the subjects which will crop up. An overview can also suggest strategies as to how different users should use the chapter or document, for example, whether to skip certain sections or whether to read another chapter first. It may also provide references to other sources of information.

Reassuring the Reader

Another way of organising information into manageable sections is to provide regular "breaks" for readers where they are given time to rest, absorb what they have just learned or even just to have a clear point in the text at which they can close the book and still feel they have achieved something worthwhile. Some sources such as Price (1984:91) suggest that such breaks may come in the form of a congratulatory remark such as "*Well done, you've just learned how to XXX. Why not take a few moments to try out*

your new skills?" or even a suggestion that they make a cup of coffee and relax for a moment. It is unlikely, however, that this would have the desired effect as readers would probably object to such a patronising and condescending tone. A more useful approach might simply be to recap on what the chapter has just covered or provide an exercise to test this new knowledge.

More difficult information requires a different type of break which is provided before the readers are actually confronted by the information. For example, if a potentially worrying message is about to appear or the chapter is about to deal with complicated information, the reader should be reassured and told not to worry. There is no harm in admitting to readers that that a particular error message may be worrying or that certain information is complicated; pretending that everything is fine and easy will serve only to make readers feel inadequate and stupid. But such admissions need to be coupled with clear instructions and reassurances that the reader will be able to manage it.

Content

Perhaps one of the most important factors in producing good user guides is honesty (Markel 2003:14). While this may seem fairly straightforward, it is surprisingly easy to mislead the reader, for example, by overstating the ease with which a particular function can be performed, by glossing over a bug in the software and hoping that users will either give up or blame themselves when things do not work as expected. Similarly, deliberately omitting information which readers need to know can have a range of consequences ranging from frustrated users to serious damage or even injury.

Following on from the notion of honesty, user guides need to be comprehensive. Just as we cannot deliberately omit information, we need to ensure that the user guide provides all of the information a user needs in order to use the product. As mentioned in our discussion of structure, this information needs to be presented when the reader needs it and not before.

One of the biggest problems, however, with user guides is accuracy and it is this, perhaps, which gives rise to a certain distrust of user guides among the general public. According to Markel:

> …the slightest inaccuracy will, at the least, confuse and annoy your readers. A major inaccuracy can be dangerous and expensive (Markel 2003:15).

There is, therefore, an ethical responsibility for technical writers and technical translators to ensure that the information contained in a user guide is easy to understand and that it is adequate for what the readers need to do. In producing user guides whether as writers or translators, our duty is to ensure that users can do what they need to do as easily and as safely as possible.

Language

The actual language used in a user guide is probably the most critical factor in determining its quality and effectiveness. Indeed, the text provides the sensory stimulus which conveys the information to the readers' brains. However, just like language itself, the factors which govern how effectively the text is used are equally vast. There are myriad guidelines, rules and regulations such as EN 62079 which are aimed at improving the standard of language but it would be impractical to discuss them all here. Rather, we can group guidelines into a number of general areas.

Clarity and Word Choice

Reminiscent of the old adage "less is more", a commonly held tenet of technical writing is that texts should be as brief and concise as possible and writers (not to mention translators) should eschew verbosity (D'Agenais & Carruthers 1985:100-101; Weiss 1985:148-9, 152). According to Weiss (1985:148) the most frequent "offenders" with regard to verbosity are what he calls "smothered verbs". A smothered verb, also known as a nominalization, is a verb that has been converted into a noun, e.g. "they conducted an investigation" instead of "they investigated". Nominalisations involve using a phrase where a single word would have sufficed and also encourage the use of unwieldy passive constructions.

Conversely, however, it is possible to be overly concise and compress text to such an extent that it becomes incomprehensible or ambiguous. The notion that text can become ambiguous as a result of excessive compression is echoed by Ramey (1989) who describes the incidence of Escher effects in texts. Escher effects - named after Escher's famous two faces / one glass picture – result in a phrase or piece of text having two or more possible meanings and force readers to truly study the text in order to ascertain or deduce which meaning of the text is the intended one.

The following examples illustrate Escher effects in text:

- input mode
- operating system file specification rules
- programming error messages

Each of these examples can have a number of possible interpretations. Taking the first example we can see that "input" can be read either as a verb or as a noun. So it is conceivable that one reader will regard "input mode" as a command – that the reader is required to input or specify the mode. Meanwhile, another reader may regard "input mode" as a state where "input" modifies or qualifies "mode".

The specific type of words used in a text can play an important role in its quality. D'Agenais & Carruthers (1985:106) suggest that positive words be used instead of negative words because, presumably, negative words have an undesirable effect on readers. The authors give the following example which is admittedly a little contrived but which does illustrate the point:

- Lock the door when you leave.
- Don't neglect to lock the door when you leave.

D'Agenais & Carruthers (*ibid.*) go on to say that words can be used to smooth the transition from idea to idea, sentence to sentence and paragraph to paragraph. The purpose of this is to avoid abrupt changes which can leave readers wondering where to go next. This idea is consistent with the theory behind the Müller-Lyer Illusion (Coe 1996:29). Figure 4 shows two lines, A and B. Both of these lines are of equal length, and each has arrowhead tails: on line A they point back over the line and on line B they point away from the line.

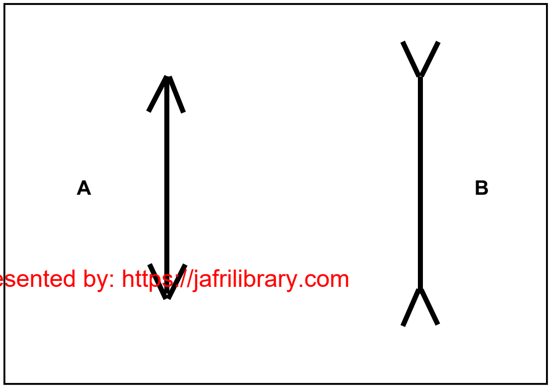

Figure 4: Müller-Lyer Illusion

Although both of these lines are the same length, the way our brains perceive the lines tricks us into thinking that line B is longer than line A. The reason for this is that the arrowhead tails on line A direct our attention back onto the line while the arrowhead tails on line B direct our attention away from the line. Similarly, textual transitions provide a link between ideas or stretches of text and offer a pointer directing the reader where to go next (Coe 1996:29).

Returning to more general aspects of word choice, it is, perhaps, useful to remember that a key goal of user guides is that they should present information in a simple manner. Simplicity of language can be obscured by a number of word choice factors: jargon, euphemisms, neologisms and abbreviations / acronyms.

Jargon

Each and every discipline, be it biology, precision engineering, electronics or meteorology has its own vocabulary of specialised terminology. This terminology is frequently referred to as jargon (White 1996:191; Mancuso 1990:186). Indeed, specialised terminology is essential in order to avoid ambiguity and to accurately communicate ideas and concepts. However, this terminology can also be an irritation and hindrance when misused (White 1996:192). The problem is, according to Mancuso (*ibid.*) that "experts use too much jargon in documents meant for less well informed audiences". The general consensus is that jargon should be used in a way that is appropriate to the abilities and level of knowledge of the audience (Mancuso 1990:186-7; White 1996:192). Where it is essential or unavoidable that jargon be used, the specialised terms should be properly defined (Mancuso 1990:186).

Euphemisms

Euphemisms are figures of speech which are used to describe things using milder, less unpleasant terms. They are generally used to soften or lessen the impact of harsh or unpleasant words or ideas. Euphemisms are frequently longer words or phrases and their meaning or relation to the actual object or action being referred to is less than obvious.

The problem with euphemisms is that while they are often quite clever, creative, linguistically interesting and occasionally amusing, they obscure meaning, confuse readers and generally make the text less accessible. In addition, because of their size, they make the text longer and more cluttered (Mancuso 1990:191).

Neologisms

Neologisms are, according to Mancuso (1990:197), the work of "arrogant" authors who like to create new words. Mancuso continues by saying that these newly created words are generally only understood by the author and a few others and they confound most readers. Admittedly, such a view is quite extreme and occasionally neologisms are necessary; they should, however, be used sparingly.

Acronyms and Abbreviations

Acronyms and abbreviations can affect the clarity and accessibility of a text in much the same way as jargon. Although many computer-related acronyms and abbreviations are becoming more widely known than they used to be (Mancuso *ibid.*), many are not yet in common usage. Thus, according to D'Agenais & Carruthers (1985:109), those that are not commonplace and understood by everyone should be explained. A popular way of dealing with acronyms and abbreviations is to use a glossary which explains them (Mancuso 1990:197; D'Agenais & Carruthers 1985:109).

Of course, clarity can also be affected by the ambiguous use of "ordinary" words. Ambiguity usually arises, according to White (1996:190), as a result of one or more of the following problems:

Improper word choice

- Using ambiguous words which can have more than one meaning in a particular context.

Unclear pronoun reference

- Pronouns must co-refer only with the noun phrase intended by the writer.

Squinting modification

- Sometimes a word can modify the phrase preceding it to give one meaning but also modify the phrase following it to give a different meaning.

Ambiguous relationships

- Using co-ordinating conjunctions such as "and" when a subordinate relationship is intended.

Sentences

> If words represent the colors of the palette available to the writer, sentences are the lines that create shapes in a composition (Houghton-Alico 1985:54).

Having looked at a number of factors relating to word choice and clarity, the next logical step is to look at how sentences affect the quality of a user guide. In line with our previous discussion of why texts should be concise, the issue of repetition and redundancy is worth examining. Firstly, we need to distinguish between repetition and redundancy. Repetition involves repeating words and phrases throughout a document in order to reinforce information, reiterate product benefits or to get readers to do or remember something. There is a definite purpose to repetition – perhaps merely to assist in the habit formation process (Raskin 2000:18-21). Redundancy, on the other hand, is "stated or implied repetition with no purpose" (Mancuso 1990:202). Redundancy can take the form of superfluous adverbs, hedge words, unnecessary emphasis or repeating information in a different form.

The flow of information in sentences is also of great importance with regard to the readability of the text. Indeed, Weiss (1985:150) argues that "the secret of the readable sentence is that the 'payload' of the sentence [...] is at the end". The payload is essentially the most important part or "nugget" of information the author wants to convey using the sentence. The reason why the payload should be at the end is, according to Weiss (*ibid.*),

that the last part of the sentence is the best remembered by readers. Similarly, in the case of instructions, a cause-effect format should be adopted (SAP 1997:4ff). Accordingly we would, for example, rewrite the following sentence:

The tab marked Properties allows users to configure the modem's settings.

as

To configure the modem settings, click the Properties tab.

Parallelism

Parallelism is a phenomenon which is widely recognised as a fundamental issue in sentence structure (D'Agenais & Carruthers 1985:104; Mancuso 1990:231; White 1996:182). Essentially, parallelism means that parts of a sentence which are similar, or parallel, in meaning should be parallel in structure. Parallel constructions can also be described as instances where two or more groups of words share the same pattern (White 1996:182). Thus, we can see that parallelism can occur on both a sentence level and on a sub-sentence level. The following sentences illustrate parallelism.

If you want to open a file, click Open.

If you want to close a file, click Close.

Parallelism can also occur in lists as shown below:

To connect to the Internet you will need:

- a modem to connect to your PC
- drivers for your modem
- a telephone line
- a dial-up account from an ISP

When there is a lack of parallelism, some of the grammatical elements in a sentence do not balance with the other elements in the sentence or another sentence. What makes this undesirable, apart from potential grammatical errors, is that it distracts the reader and prevents the message from being read quickly and clearly (Mancuso 1990:232).

Returning to the examples of parallel constructions given above, we can illustrate how a lack of parallelism can affect the clarity and readability of a section of text. What were once clear sentences, become the following confusing examples:

If you want to open a file, click Open.

The Close button should be pressed to close a file.

To connect to the Internet you will need:

- a modem to connect to your PC
- drivers for your modem
- a telephone line must be available.
- also, contact an ISP to get a dial-up account

Parallelism is not just important in avoiding grammatical and comprehension problems, it is also very useful in reinforcing ideas and learning. The grammatical symmetry of parallelisms helps readers remember information more easily (White 1996:183). The notion of parallelism is closely related to that of Iconic Linkage which we will discuss later in Chapter 4.

Style

When we speak of style, we really mean the overall tone of the text and how authors express themselves – essentially, how the authors relate to their readers. It is generally acknowledged that a conversational style of writing is the best approach when producing and translating user guides. Mancuso (*ibid.*) ventures by way of an explanation, that the way we normally write is generally unsuitable for explaining ideas. When we explain ideas orally, we are concise, to the point and we avoid awkward or complicated constructions. Indeed, D'Agenais & Carruthers (*ibid.*) maintain that most people communicate better when they are speaking than when they are writing. A possible reason for this is, according to the authors, that people tend to "write to impress rather than to express". Mancuso and D'Agenais & Carruthers agree that the best way of avoiding "stilted prose, using passive voice verbs and other awkward constructions" (Mancuso 1990:149) is to explain things orally or in speaking mode rather than in writing mode.

Using a conversational style does not, however, give authors free rein to use slang, to be excessively informal or to be imprecise or ambiguous. While oral communication has the benefit of instant feedback from the receiver's reactions, written communication does not have this aid and so the potential for misunderstanding must be minimised.

Verbs

Verbs are the engines of sentences – they make the sentences meaningful and make a text more than just a list of words. The way in which verbs are used affects the way the text works and how easily the reader assimilates information. We can categorise our examination of verbs as follows:

- Strong / weak verbs
- Active / passive voice

- Imperatives
- Compound verbs

Strong and Weak Verbs

The differentiation between strong and weak verbs can be quite subjective and is rather elusive. It would, perhaps, be easier to define the two terms using a number of examples of strong and weak verbs. Mancuso (1990:174) suggests that strong verbs might include *weld*, *singe*, *salivate*, *bulldoze* and *inject*. Weak verbs, he continues, include the various forms of the verb *to be* and the verbs *do*, *make*, *provide* and *include*. Strong verbs, he maintains create images; they add a sense of action to a text. On the other hand, weak verbs say little, if anything and result in the reader having to spend more time "deciphering meaning rather than reading it" (*ibid.*).

From the examples given below, we can see that strong verbs are those that actually reflect the function or action in question. The following sentence is rewritten to illustrate examples of strong and weak verbs:

The function of the hard disk is to allow you to store data.

The hard disk stores data.

The benefit of using strong verbs is that it allows readers to understand information more quickly. Additionally, as can be seen in above example, strong verbs allow for more concise constructions.

Nominalisations, i.e. verbs that have been transformed into nouns, are just as unhelpful as weak verbs in that they obscure meaning and add to the workload of readers. An example of this would be as follows:

The setup program results in an update of the registry.

If we remove the nominalization, we get the following:

```
The setup program updates the registry.
```

Active and Passive Voice

The terms "active" and "passive" voice are old metaphors for certain grammatical constructions. Active voice constructions contain subjects that "do something". These constructions have positive connotations of action, dynamism, energy and determination (White 1996:181). Passive voice constructions, on the other hand, contain subjects that do not do anything. These constructions have the opposite connotations to active voice constructions. The passive voice is typified by the following characteristics:

- The subject is acted upon.
- The predicate generally contains an auxiliary verb that is in the form of to be.
- The sentence contains a prepositional phrase

While it may be helpful to switch between active and passive voice in order to emphasise either the subject or the logical object (White 1996:182), it is widely held that the passive voice interferes with the clarity of sentences (White *ibid.*; Mancuso 1990:156-171; D'Agenais & Carruthers 1985:102-3). It is also more difficult for readers to understand the sentence because of problems with identifying the actor and also because of delayed meaning (Mancuso 1990:166-7).

Imperatives

Using the active voice in conjunction with the imperative mood is an important strategy in procedural texts where the reader is required to either perform certain tasks or refrain from carrying out certain actions. In contrast to constructions that do not use the imperative, there is no confusion as to who is to carry out the task because the second person pronoun *you* is implicit (Price 1984:103). Take, for example, the following sentence:

```
The necessary drivers must be installed on the PC.
```

From this sentence it is not clear who is supposed to perform the task. Is it the reader or is it someone from the IT department? It would be better to phrase the sentence as follows:

Install the necessary drivers on the PC.

Assessing User Guide Quality

The evaluation of user guides takes place on a number of levels. While it would be tempting just to test a user guide to see that it accurately reflects the software, such an approach would provide only a partial picture. As we have seen in preceding sections, the design and layout of a user guide as well as the way it is written are important factors in producing a user guide. In the following sections we will examine several methods for assessing user guides.

Readability

In addition to ensuring that writing is clear, consistent and concise, readability testing also indicates whether the text is at the correct level for the intended audience. There are numerous methods for measuring readability including the *Flesch Readability Test*, the *Lensear Write Formula*, the *Fog Index*, *Fry's Readability Graph* and the *Clear River Test*.

Most of these methods involve selecting a sample of text between 100-200 words in length. Each of the methods mentioned above express readability in terms of the proportion of various features such as syllables, monosyllabic words etc., average sentence length etc. Methods such as the Fog Index regard words with more than three syllables as "difficult" while words with less than three syllables are regarded as "easy".

The Flesch Readability Test developed by Rudolph Flesch examines readability as a relationship between the average sentence length and the average word length; the shorter the sentence and the shorter the words, the more readable the text. The readability of a text is presented using a scale of 0 to 100; the higher the score, the easier the text is to read (D'Agenais & Carruthers 1985:113).

The Fog Index identifies easy words in a text, i.e. words with one or two syllables, and calculates readability as a function of the average sentence length and the percentage of "hard" words, i.e. three or more syllables.

Readability is then expressed in terms of the number of years schooling needed to read the text with ease. Similarly, the Lensear Write Formula calculates readability on the basis of the proportion of monosyllabic words in a text using the following formula:

$$\textbf{Lensear Score} = \frac{\textbf{No. of monosyllabic words + (3 x no. of sentences)}}{\textbf{Total Words}} \quad \textbf{x 100}$$

Figure 5: Lensear Write Formula

Using the Lensear Write Formula, a score of 70-80 points is regarded as ideal for user guides (D'Agenais & Carruthers 1985:114).

Like the Fog Index, Fry's Readability Graph expresses readability in terms of the level of schooling needed in order to read a text with ease. Taking a sample of 100 words, it calculates the average number of syllables per sentence and the number of sentences per 100 words before expressing readability as a function of the two.

Fog Index = (Average Sentence Length + Percentage of Hard Words) x 0.4

Figure 6: Fog Index Formula

The Clear River Test combines several of the features of the preceding methods and analyses readability in terms of the number of words per sentence, per paragraph, per punctuational pause and the number of syllables per 100 words.

These tests can prove very useful in providing an overview of how effective a user guide is in terms of readability but they do not explain why a user guide is ineffective despite being readable. Is the text poor because of the register used? Does it contain too much jargon? Are concepts not explained clearly? Does the text contain ungrammatical constructions? Indeed, George Klare, a leading academic in the field of readability evaluation formulae concedes that readability assessments are of limited use in assessing computer documentation and that in some cases, such methods were not even designed for use on such texts (Klare 2000:2-3). It is clear that in order to pinpoint precisely what errors contribute to a text's underperformance, we need to find a more comprehensive evaluation method.

Usability

Another approach to determining the effectiveness of a user guide is to establish how effective it is in achieving its purpose and how easy it is to use. In contrast to readability assessment methods which examine linguistic and technical features from the point of view of the text, usability introduces a new element into the equation, i.e. users. Usability assessment evaluates linguistic and technical features such as those described in previous sections and assesses the sum total of all of their contributions from the point of view of the user. Instead of considering only the readability of text or whether the style is appropriate, usability is concerned with the ease with which users (readers) can access and assimilate information and then use it to complete their intended tasks, i.e. use the software.

A simple way of testing the usability of a user guide is to gather a group of people who reflect the actual audience for the user guide and have this group use the software on the basis of the user guide. The purpose here is to see where the readers succeed and where they go wrong, where they have difficulties and where they need more help. Usability testing of this type (as defined in technical writing literature) tests the user guide for logic gaps and inadequate clarity. It determines whether readers can actually use the user guide effectively and efficiently and whether users actually learn from it.

Usability is a central element of what is known as Human-Computer Interaction (HCI). This area is concerned with examining the interactions between humans and computer systems (e.g. software). The description of usability provided in the preceding paragraphs is admittedly rather simplified and rudimentary. In order to fully understand usability, it is necessary to understand the primary component of these interactions, i.e. humans. In the following chapter we will examine usability from the point of view of users and discover the mechanisms that must be understood and accommodated to ensure usability.

Conclusions

This chapter introduced the genre of software user guides and placed it within the overall context of technical communication. In doing so it provides translators with an insight into the various features and characteristics of these texts as well as the expectations typical readers have of user guides. By understanding these issues, translators can gain a better understanding of

what it is they need to achieve in order to produce effective translations which will be accepted by the target audience.

It is clear from this chapter that user guides are just one product of technical communication, yet they are, perhaps, one of the most visible products. It could be argued that they are one of the crucial types of technical document because they are instrumental in allowing new users to learn how to use new software. We have seen that the perceived ease of learning as facilitated by user guides can be a decisive commercial factor for software products.

This chapter also examined in detail the components of user guides. From this it emerged that a "good" user guide is more than just a collection of clearly phrased instructions or a repository of all information relating to a piece of software. Rather, user guides should ideally contain information that is targeted at the needs of the users and what they want to do. They present users with the knowledge they need in order to perform a task, when they need to perform it. Information is "fed" to users in a measured, logical and timely way.

Beyond the purely stylistic and content-related issues, a range of other factors such as layout, typography, presentation, structure etc. influence the effectiveness of user guides. All of these factors paint a more holistic picture of the nature of user guides than that which is reflected in the methods commonly used to assess the quality of user guides. Readability tests such as the *Flesch Readability Test*, the *Fog Index* or the *Clear River Test* coupled with technical accuracy checks merely assess a small part of user guides.

Line spacing, white space and information chunking point to some form of understanding of how humans read text and perceive information. Simplicity of language, clear instructions, the use of parallel structures and active verbs and the avoidance of euphemisms etc., all draw on characteristics of the way humans decode, understand and absorb information. It is clear, therefore, that we need to examine these factors from the point of view of the person reading the user guide.

One area raised in the chapter which will be discussed in much more detail later on is that of usability and usability testing. Unlike readability testing, usability testing seeks to understand all of the factors that influence how well users can use a user guide, whether it is an original source language text or a translation. Usability testing adopts a suitably broad approach which centres on the reader and it will give us a deeper understanding of why some user guides are easier to use than others. The following chapter will begin by defining usability and discussing its importance for users. We will then look at the processes and systems that are called into play when

we read a user guide. By understanding readers' cognitive abilities, preferences and limitations, we can begin to identify those aspects of user guides that facilitate the transfer and assimilation of knowledge necessary to use software – the stated purpose of user guides. Thus, any discussion of usability requires a thorough understanding of the human cognitive system and its processes.

Chapter 3
Understanding Users

If usability is the measure of how easily and effectively people can use something (in this case, a user guide), a key component in understanding it is studying the interaction between humans and user guides. Since usability refers to the extent to which people find something easy to use, to understand usability we must first understand the users who are the ultimate judges of usability. In this chapter we will examine ways of modelling the human cognitive system and discuss how it works. Human sensory, cognitive and learning processes will be examined as well as the way we remember information. The various processes involved in assimilating and interpreting information will also be explained.

Usability

> A computer shall not waste your time or require you to do more work than is strictly necessary" (Raskin 2000:6)

We have previously discussed how making user guides more accessible and usable is a primary objective for technical communicators. It would be easy to produce a simple working definition of usability such as "ease of use". However, such a definition by no means explains the true nature of usability and the factors affecting it.

Another common fallacy is to confuse usability and usefulness. While ostensibly related, they are poles apart in terms of their relationship to products. Usefulness refers to the potential uses users can find for something whereas usability refers to how well users can use it (Landauer 1995:4; Ehn & Löwgren 1997:301; Dumas & Redish 1993:5).

However, defining usability as a measure of how well users can use something is a slight over-simplification. In the ISO 9241-11 standard

"Ergonomic requirements for office work with visual display terminals", usability is defined as:

> The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context.

In other words, usability refers to how well a given user will perform a given task at a given time. There are difficulties, however, in this definition with regard to the phrase "specified users". A number of theorists maintain that attempting to define the concept of a user is highly problematic and possibly even futile (Ehn & Löwgren 1997:299; Bannon 1991:26-27). This is because there may be a vast number of different people, all with different backgrounds, knowledge, abilities, skills and patterns of use who will all be using the product. It would be prohibitively difficult to profile each type of user in terms of the aforementioned criteria. In view of this, Ehn & Löwgren (*ibid.*) propose that our focus should be on the situation of use, i.e. where, when and how the user uses the product.

Dumas & Redish provide a definition of usability which is less specific than the ISO 9241 definition given above but which nonetheless provides additional insight. According to their definition, "usability means that the *people who use the product* can do so *quickly and easily* to accomplish *their own tasks* [emphasis in original] (1999:4). Here, the crucial factor is the fact that users are using the product to perform *another* task. The use of the product is secondary to a user's true intention. We can see, therefore, that usability does not depend on the product *per se*, but rather on the people who use it. A usable product is one which is appropriate to the **tasks** users want to carry out. Indeed, according to Faulkner (1998:7) "the very best systems and the very best interfaces will be overlooked entirely by the user" and ideally, all the user should see is the task and not the system.

Dumas & Redish (1999:4-6) examine the relationship between usability and users under the following headings:

- Usability means focussing on users
- People use products to be productive
- Users are busy
- Users decide how usable a product is

Usability Means Focussing On Users

In order to make a usable product it is vital to understand real users. People such as managers and developers do not represent real users.

People Use Products to be Productive

Software products are tools which people use in order to do something else. People judge the product on the basis of the time it takes them to do something, the number of steps they have to perform and how successful they are in doing the task. The aim is to make products so easy to use that users can perform their tasks more quickly.

Users Are Busy

The usability of products is gauged by users in terms of how quickly they can get the product to do something. A product may have precisely the functionality a user needs to perform a task but if the function cannot be accessed or used within the time the user is prepared to devote to the task, it will be useless. This idea is virtually identical to Landauer's distinction between usefulness and usability discussed above.

Users Decide How Usable a Product Is

Regardless of how well developers, managers or marketing people think something is designed, the ultimate judge of usability is the users themselves. If the effort needed to perform a task outweighs the benefit, users will regard the product as unusable.

The Scope of Usability

In previous paragraphs we referred to usability in terms of its relationship to "products" or as "systems" and the product's relationship to users. When we speak of users and usability we are actually referring to the interactions between users and products or systems. This is indeed convenient but when we speak of products or systems we are referring to a collective of various different components all of which make up the whole that is the software system. Such components include hardware, software, menus, icons, messages, user guides, quick reference guides, online help and training. All of these have a bearing on usability and conversely, the usability of each of these factors affects the usability of the system as a whole. This synergy between the components is echoed by Dumas & Redish (1999:6) who state that "changes in technology have blurred the lines among these product pieces, so that it is no longer useful to think of them as separate entities". As

such, a user guide that is less than satisfactory from a user's point of view will adversely affect the overall usability of the system because user guides form a core part of the system. In the following section, we will examine users from the point of view of those cognitive processes and abilities which affect the way users use user guides.

The Human Cognitive System

If we take a basic view, computers are information processing systems. Information, or data, is manipulated, created, modified, accessed and stored. Similarly, the human mind can also be regarded as an information processing system. As such, the broad model of a computer can be used as an analogy for describing the human information processor (Card *et al.* 1983:24; Downton 1991:20). We can draw several comparisons between the two contexts in that they can both be said to consist of memory, processors, interconnections, rules etc. However, such an approach can only be used for illustrative purposes as the structure of a computer does not necessarily reflect the structure of the brain. Indeed, there is still some debate about whether certain components of the mind are distinct physical locations or merely different functions of the same physical location (Card *et al.* 1983:23,36; Dix 1998:27; Faulkner 1998:33-34). Raskin (2000:12) warns against using current technology as figurative models because such models rapidly become outdated and quaint. Nevertheless, using computers as an illustrative model allows us to conveniently examine the human mind as a series of subsystems. If we return to the idea of a computer we can see on a very basic level that:

- information is input into the computer
- the information is processed, and
- an appropriate response or output is prepared

Applying this scheme to humans we can divide the human mind into the following subsystems (see Card *et al.* 1983:24; Downton 1991:20):

- the perceptual/sensory system
- the cognitive system
- the motor system

For our purposes here it is convenient to discuss the perceptual and motor systems together as they are similar to the basic notion of a computer's input/output system. This model, however, omits a fundamental factor common to both computers and humans: information is stored and accessed. And so, to make the model more accurate in terms of functions we need to incorporate memory into it. We can use the following components to examine the human system (see Dix 1998:12):

- input/output

- memory
- processing

The Human Input / Output System

As already stated, humans interact with the outside world and with computers through the exchange of information. This exchange relies on information being sent and received; in other words the input and output of information.

Information input for humans takes place through the five senses: sight, hearing, touch, taste and smell. For most people, the first three senses are the most important, especially in terms of human-computer interaction (Faulkner 1998:13; Dix 1998:13). Even though the senses of taste and smell are valuable senses for humans, it is not clear how they could be utilised in our interactions with software or documentation (Dix *ibid.*) and they will not be discussed further here. Similarly, the senses of hearing and touch, although invaluable for humans, are of little relevance in our examination of printed user guides; these senses will not be discussed here either.

Sight – The Visual Channel

The way in which humans see relies on a highly complex system which functions in two stages: the physical reception of visual stimuli and the processing of these stimuli (Dix 1998:14). While the physical limitations of the eye mean that we cannot see certain things (e.g. ultraviolet or infrared radiation etc.) the processing abilities of humans means that we can organise visual perception in terms of motion, size, shape, distance, relative position and texture (Downton 1991:14), even if some information is missing or incomplete (Dix *ibid.*).

The Eye

The basic principle underlying human sight is the reception of light reflected from the physical world. This light represents a stimulus which is then interpreted. At the core of this process is the eye which is a firm, semi-solid orb consisting of several layers (see Figure 1).

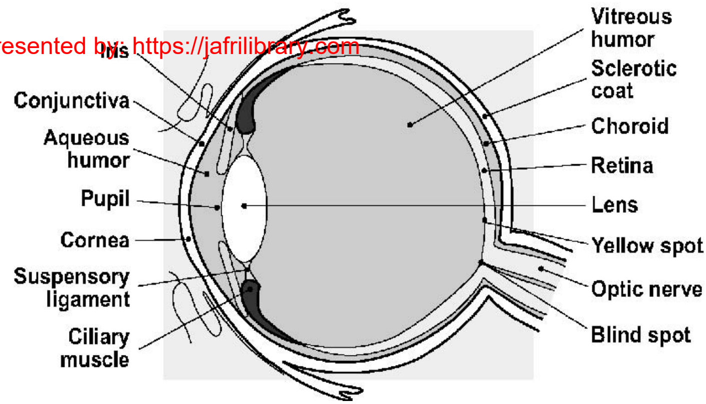

Figure 1: The Structure of the Eye

The first of these layers is a white outer layer called the sclera which not only maintains the eyeball's shape but provides both protection and a means for attaching the muscles that move the eyeball. Inside the sclera is the choroid, which is a dark layer containing many blood vessels which nourish the various structures of the eye. In addition to this, the dark colour of the choroid layer prevents light from being reflected internally within the eye.

The innermost layer of the eye is the retina which contains cells which are sensitive to light, i.e. photoreceptors. Towards the front of the eye the choroid layer becomes thicker and forms the ciliary body. This part of the eye contains muscles called the ciliary muscles which are used to control the shape of the lens which is suspended by suspensory ligaments. Another part of the choroid layer is the iris, a coloured circular body which is similar to the shutter on a camera. The function of the iris is to regulate the amount of light entering the eye. Inside the eye, we can distinguish between two different chambers, one in the main body of the eye and the other between the cornea and the lens. The first chamber is filled with a clear glutinous substance called the vitreous humour. The second chamber contains a less

viscous clear fluid called the aqueous humour. Both humours contribute to maintaining the shape of the eyeball.

The purpose of the eye is to receive light and convert it into electrical impulses which are then sent to the brain via the optic nerve. When light enters the eye, it is refracted by the lens and focused to form an image on the cells on the back of the eye (see Figure 2). These cells make up the retina.

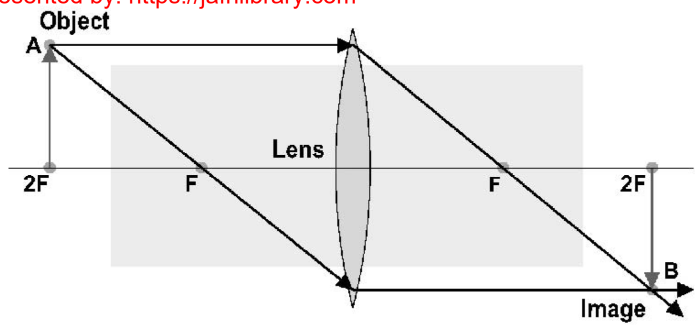

Figure 2: Refraction of Light Through a Lens

In a process called accommodation, the shape of the lens is altered by the ciliary muscles to focus images from distant objects: when viewing distant objects, the muscles relax, making the lens thicker. When viewing close objects, the muscles contract to make the lens flat and thin.

When light enters the eye it has two main properties which are of use to the eye: brightness (intensity) and colour (wavelength) (Faulkner 1998:14). The light is focused onto the retina which contains two types of photoreceptor cells (i.e., cells which are sensitive to light): *rods* and *cones*.

Rods are very sensitive to light and it is these cells which allow us to see in low-light conditions. The average eye has around 130 million rods (*ibid.*), most of which are located towards the edges of the retina (Dix 1998:14). Rods, however, are not very good at detecting detail and are subject to saturation. Saturation essentially means that the cells become “overloaded” by light and this phenomenon explains the temporary blindness experienced when moving from dark into light (*ibid.*). Rods are also unable to detect colour.

The second type of photoreceptor is the *cone*. There are three types of cone, each being sensitive to a different wavelength of light. In the average eye there are approximately 6-7 million cones and they are mostly concentrated on the fovea which is the target on the retina where images are focused. Cones are less sensitive to light and they can detect colour. This means that we can detect colour if the object is in our main field of vision as opposed to our peripheral vision. There is a point on the retina which does not have any photoreceptors and which cannot detect light. This is called the *blind spot* and it is where the optic nerve connects to the eye.

In addition to photoreceptors, the retina also contains special nerve cells called *ganglion cells* of which there are two types: *X-cells* and *Y-cells*. X-cells are concentrated in the fovea and are responsible for quickly detecting patterns. Y-cells are distributed across the retina and are responsible for detecting movement. But how is it that these physical structures transform the physical stimulus of light into useful information which we can understand?

Visual Perception

As mentioned above, only two physical properties of light are detected by the eye: intensity and wavelength. This "raw information" must be filtered and processed before it becomes meaningful input for the brain.

The human eye is sensitive to light with a wavelength between 400nm and 700nm (Downton 1991:18). This wavelength range is known as visible light. We know that the eye is sensitive to the wavelength and the intensity (see page 103) of light so how do we distinguish between the approximately 128-150 distinct colours and 8 million different shades? (Downton 1991:18; Dix 1998:18). To understand this we need to examine the components of colour, namely, *hue*, *intensity* and *saturation*.

Hue corresponds to a superordinate class of colour, i.e. it is the same as what we mean when we say "colour" in everyday conversation. Hue is determined by the wavelength of the light. So for example, blue light has a wavelength of approximately 400nm and red light has a wavelength of approximately 700nm. Intensity is the brightness of the colour and saturation is the amount of white in the colour. By varying these two components, it is possible to detect between 7 and 8 million different shades of colour. However, the number of colours which can be realistically detected by an untrained eye in isolation is in the region of 7-10 (Dix 1998:18; Downton 1991:18).

Brightness is a subjective response to light. There are no absolute levels of brightness and our perception of it is determined by our own personal preferences, physical make-up and physiological state (e.g. migraine headaches make us more sensitive to light). Brightness is affected by *luminance* which is the amount of light emitted or reflected from the surface of an object. Unlike brightness, luminance is a physical characteristic and can be measured using a photometer (Dix 1998:18) and is quantified in *candelas per square meter* (Faulkner 1998:19). In practical terms, this means that a brighter object (e.g. text) will be easier to see and read. However, increased brightness also makes the eye more susceptible to flicker (Faulkner 1998:20).

Contrast is related to luminance in that it is the relationship between light emitted by an object and the light emitted by the background. This is illustrated in Figure 3.

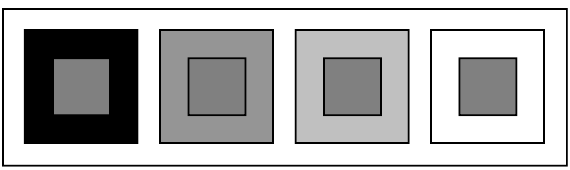

Figure 3: Brightness contrast

In this diagram, the small squares are all equally as bright as each other but because of the contrast with each of the backgrounds, the squares look brighter or darker depending on how dark the background is.

Saturation is the amount of white light in a colour. But to fully understand saturation, we must first look at the difference between *chromatic* and *achromatic* colours. Chromatic colours such as blue, red, green, etc. have hue and are affected by the wavelength of light. Achromatic colours, on the other hand, such as black, white and neutral grey do not have a hue and are not affected by wavelength (Faulkner 1998:17).

Saturation refers to the extent to which a colour is chromatic or achromatic. The more achromatic colours go to making up a colour, the less saturated it becomes. A colour's saturation is said to be zero when it is entirely grey.

The Boundaries of Visual Processing

The basic principle underlying human sight is the reception of light reflected from the physical world. But this is only a small part of the visual channel. This light represents a stimulus which must then be interpreted. What makes this channel so valuable is what we do with the information we perceive. This information must be processed and transformed so that we can form an interpretation of the images we see.

In processing visual information, our expectations, experience and knowledge play a key role. For example, if we know that a truck is 15 feet high and 40 feet long, we will always perceive it as such even if we view it from a distance. It is this ability which allows us to make sense of unexpected, faulty, contradictory or incomplete information and allows us to resolve ambiguous information. Our expectations in relation to the world around us are largely determined by the context. Accordingly, one set of criteria may apply in one particular situation, e.g. a truck appears huge and has a trailer and 18 wheels when viewed up close, while different criteria apply at other times, e.g. a truck appears small and has a trailer and 18 wheels when viewed from a distance.

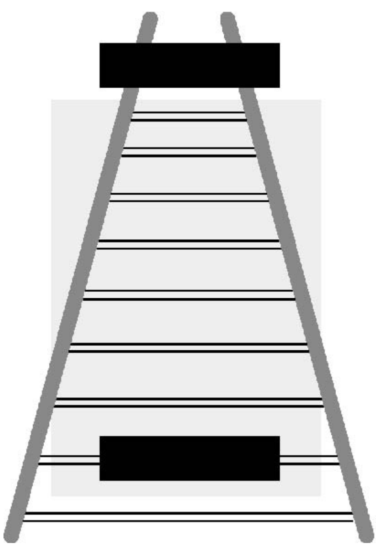

Figure 4: The Ponzo Illusion

Unfortunately, this ability to make sense of ambiguous or contradictory information is not perfect and is prone to errors and interference. This can be illustrated by optical illusions such as the Ponzo illusion (Figure 4) where the top line appears longer than the bottom line when in fact both are the same length. This can be attributed to an incorrect application of the law of constancy whereby the top line seems further away and is made to appear bigger (Dix 1998:19).

the quick brown fox
jumps over the
the lazy dog

Figure 5: The Proof-reading Illusion (Dix 1998:21)

In the proof-reading illusion, most people reading the sentence quickly will miss the second "the". However, on closer examination, people will spot the mistake. This is an example of how our expectations compensate for unexpected information when reading.

Reading

Perhaps one of the most complex applications of visual processing, and the one most closely related to translation is the ability to read. Reading as an activity consists of several stages. Firstly, the visual pattern or appearance of the word is perceived. This physical image must be decoded by matching it against our own semiotic knowledge (signs such as letters and words) and then interpreted on the basis of syntactic and semantic analyses which operate on phrases or entire sentences (Dix 1998:22).

When we read, our eyes move in a series of saccades. This means that the eyes do not move smoothly but rather in a stop-start manner. Each saccade consists of a brief period of motion followed by a fixation (Card *et al.* 1983:50) which is when the eye is at rest and when perception occurs. Fixations account for 94% of time spent actively reading. The eye moves both forwards and backwards to read and re-read text. These backward movements are known as regressions and are more frequent when reading complicated or difficult texts.

Generally speaking, the average adult can read approximately 250 words per minute. This speed means that it is unlikely that each letter is scanned and decoded in series. Indeed, according to Dix (1998:22) we can recognise certain familiar words by their shape just as quickly as we can recognise a single letter. One interesting effect of this is that it is very easy to destroy the visual clues which make whole words recognisable by shape. So for instance, if we were to capitalise a word, we undo the familiarity of the word's shape and consequently the word will have to be scanned and processed as a string of letters rather than as a single meaningful unit (*ibid.*). Take for example the word "intermediate". Written like this we can recognise it almost instantly. But if we write it in uppercase like this INTERMEDIATE, it is not so immediately recognisable.

Human Output

Taking a simplified view of the human cognitive system, we can say that information is received by the sensory organs and sent to the cognitive system for processing. Once the information has been processed, a response is produced. The brain sends the necessary impulses to the appropriate part(s) of the body in order to effect this response.

Our bodies can respond physically using our hands, fingers, thumbs, feet and voice. As with many functions and activities related to humans, the effectiveness and speed with which we respond physically varies from person to person as a result of factors such as age, fitness, health or alertness. The speed with which we react to a stimulus also depends on which sense receives the stimulus: we can react to an auditory stimulus in approximately 150ms; to a visual stimulus in 200ms and to pain in 700ms (Dix 1998:26). Reaction times are not the only factors affecting human output. The actual output rate varies depending on the part of the body used to respond to a stimulus. For instance, if we use one finger on each hand in an alternating manner, we can achieve somewhere in the region of 1000 key presses per minute. If we use just one finger, this figure is around 400 key presses per minute. Vocal output allows us to achieve an output of between 180-400 words per minute (Downton 1991:26).

Perception

In the preceding paragraphs we examined the sense of sight. This is the most important sense in terms of how we use user guides. Now we will look at what we do with the information we gather from our surroundings.

Perception is more that just seeing or hearing. Perception is a complex and active process which allows us to interpret information. By interpreting the raw information provided by our sensory organs we prepare it for further processing in the cognitive system. If it were not for perception, we would simply be receivers of sensory information but we would not be able to use this information for anything. Think of a motion detector – it can detect an intruder but unless it is connected to an alarm system it cannot activate a siren or alert anyone. Of course, if an alarm system had the cognitive processing abilities of humans it would also be able to distinguish between intruders and friends. Conversely, without the sensor, the alarm system is deaf and blind – it simply cannot do anything because it receives no information.

Sensory Data Filters

With our powerful sensory systems, humans are under a constant barrage of sensory information. We receive enormous amounts of information through our eyes, ears, sense of touch etc. But it would be impossible for us to process all of this information or even be consciously aware of it all (Coe 1996:10). Indeed, we are only aware of a fraction of the sensory information we receive.

This is not a coincidence, for if we were to attempt to process everything we would waste valuable processing resources on things other than those we want to concentrate on. It is possible that such a volume of information could even overload our processing systems with less than desirable consequences. We must, therefore, organise and limit the sensory input so that we can process information in a structured and manageable way. This is done in a number of ways that make use of *thresholds*, the *cocktail-party effect* and *sensory adaptation*.

Thresholds

There are two types of threshold that we use to separate and organise sensory input: *absolute* and *just noticeable difference* (JND).

An *absolute threshold* is the smallest amount of stimulus we can detect 50% of the time. This type of threshold is largely dependent on the individual and each individual's psychological state. For instance, a user's motivation, expectations and experience are crucial in determining absolute thresholds when, for example, learning to use a new software application. Consequently, absolute thresholds are variable – the exact same stimulus may induce different responses under different circumstances and at different times (the time of day, whether the user is in a good mood etc.).

Just noticeable difference is the smallest difference noticeable between two stimuli 50% of the time. By way of example, let us consider a glass of water. Imagine we are gradually adding tiny amounts of salt to the water. We will not detect the presence of the salt at first but if we continue adding tiny amounts of salt, we will eventually begin to taste it in the water. The difference between the point where we first detect the taste of salt and the last point where we did not taste the salt is the just noticeable difference.

Cocktail-Party Effect

The cocktail-party effect allows us to filter out information which is important or relevant and separate it from the deluge of sensory information we constantly receive. The effect allows us to focus in on important information while ignoring irrelevant information. The analogy comes from the notion of a cocktail-party where many different conversations are taking place. Amidst this bustle of information and conversation, we will hear someone mention our name over the noise in the room. Similarly, we can generally choose to listen to one particular conversation and effectively "fade out" the other conversations and "turn up" the conversation we want to listen to (Preece 1994:100; Coe 1996:12).

Sensory Adaptation

Sensory adaptation describes the phenomenon whereby we become accustomed to a set of sensory inputs. For instance, if a person is working in an office and the air conditioning is turned on, the person may be distracted by the noise of the fan. However, after a while the person becomes accustomed to the new stimulus and no longer notices it. It is not until the fan is turned off and the noise stops that the person becomes aware of it again.

Thresholds, the cocktail-party effect and sensory adaptation are all mechanisms by which we select which information to process. They allow us to optimise processing resources and concentrate on what is actually

important, relevant or of interest. (These mechanisms are also important factors in attention and cognitive processing which will be discussed later). Now we have separated the information to be processed, we can look at how this information is interpreted and prepared for cognitive processing.

Ecological and Constructivist Approaches to Perception

There are a number of different theories which seek to explain how we turn basic sensory data into meaningful interpretations. These theories can be broadly categorised into the following groups: *ecological theories* and *constructivist theories*.

The fundamental difference between these two groups is that ecological theorists maintain that perception involves a process of gathering information from our environment to help us understand our surroundings. Constructivists, on the other hand, believe that visual perception is an active process based on what we actually see as well as our own previously acquired knowledge (Preece 1994:76). Using both of these elements we then construct an interpretation of the information we receive.

Ecological Approaches

This approach states that perception is a direct process whereby we detect information rather than create or interpret it. The ecological approach is not concerned with how we understand or recognise situations or scenes but rather what we need to know about a situation and how we go about finding it in our environment. This approach involves us actively exploring our surroundings and engaging in activities that allow us to find the necessary information.

Constructivist Approaches

The constructivist approach, on the other hand, maintains that visual perception is not just a direct representation of what we see but rather a model of our surroundings which is modified, transformed, enhanced and filtered using our knowledge, experience, expectations and memories. This approach sees perception as a process whereby what we see is compared against our experience of the world and an interpretation is constructed. What is more, by comparing what we detect from our surroundings against

what we know, we can deal with a wide variety of situations and, if necessary, adapt to new situations and modify existing knowledge.

Piaget's concept of schemes (Piaget & Inhelder 1969:4; Hill 1995:15; Ginsburg & Opper 1988:20-22) is a useful tool in understanding this. When people are presented with new tasks or situations, they bring with them a set of existing ideas, methods and knowledge (known as a scheme) which they will use to tackle the task. However, if this scheme is not adequate for the task, they will modify this scheme in order to incorporate new knowledge, methods or approaches. Take, for example, the driver of a car. Driving a car requires a set of knowledge such as understanding gears, using the pedals, starting the engine, stopping distances, traffic regulations, manoeuvring the vehicle, etc. Now let us imagine that this person wants to drive an articulated truck. The knowledge of driving a car is only partly useful – the rules of the road still apply as does the knowledge of using gears. But the knowledge of manoeuvring, braking distances etc. is different for trucks and will have to be modified if the driver is to successfully drive the truck. Schemes are also referred to as perceptual sets (Coe 1996:16).

Grouping and Organising Information

In order to interpret the objects we see, we need to be able to regard them as meaningful units. Under the broad category of constructivist approaches, the Gestalt psychologists such as Koffka (1935) and Köhler (1947) developed a way of grouping or organising information so that it "means" something or forms something to which a meaning can be attributed. So rather than seeing a series of separate, individual objects in isolation, we group them into units or organised "wholes" (Coe 1996:18). This is the basis for the statement on page 108 that we can recognise certain words from their shape just as easily as from the individual letters that make up the word.

The Gestalt approach to organisation provides us with 6 basic "laws" which help us organise and interpret objects: *Proximity*, *Similarity*, *Continuity*, *Symmetry*, *Closure*, and *Common Fate*.

Proximity	If objects are near each other, the average person will tend to group them together. This law applies not only to objects such as lines or shapes but also to text, tables etc.
Similarity	If objects are similar, we will group them together as a unit.
Continuity	We are more likely to perceive smooth, continuous patterns rather than abrupt or non-continuous ones.
Symmetry	If we see an area bounded by symmetrical borders or objects, we tend to group everything together to form a symmetrical figure.
Closure	If we see an object with gaps in it, we will see past the gaps and view the object as a whole.
Common Fate	If we see objects moving or pointing in the same direction, we will group them together as a single unit because they share a "common fate".

Table 1: Gestaltist Laws for Grouping Information

Prägnantz

The law of Prägnantz (Coe 1996:23) is also called the "goodness of figures" and refers to the way humans generally opt for the simplest, most obvious interpretation of an object. This "law" illustrates how we group information and compensate for missing or faulty information to produce the most probable and likely interpretation given the context. In a way which is similar to the ideas put forward by supporters of relevance theory (Sperber & Wilson 1986; Gutt 1991), humans will opt for the interpretation which is most accessible and which requires the least processing effort.

Pattern Matching

Once we have grouped the objects we see into meaningful units, we need to recognise them in order to understand what they are. There are a number of ways in which we can recognise shapes (or patterns) and which ultimately determine whether we correctly interpret them.

Prototype Matching

This method involves us storing a general, fundamental shape or pattern against which we compare objects to find a match. Essentially, this model is a very basic stylisation which is fundamentally the same regardless of any cosmetic or superficial differences we encounter from instance to instance.

Template Matching

In contrast to prototype matching which provides us with a general outline of objects, template matching involves us storing detailed patterns of each and every variation of an object we see. So rather than having a prototype for the letter "P" which states that a "P" consists of an upright line with a loop attached to the top right, template matching means we need a model or design for each "P" we encounter.

Distinctive Features

This method involves us distinguishing objects on the basis of their distinctive feature patterns. For example, a car has four wheels while a bicycle has just two wheels. With this method, we recognise objects by analysing them and matching distinctive parts of an object as opposed to the entire object.

Memory

Having discussed the sensory system we will now continue our examination of the "infrastructure" which allows the human cognitive system to work. From a human-computer interaction (HCI) and learning point of view, we can say that the sensory system is the mechanism for receiving data to be processed while memory is the mechanism which facilitates cognition and learning. Only by understanding memory can we proceed to look at how data is processed and understand how we learn and solve problems.

Memory is fundamental to virtually every one of our actions from reading, eating and walking to writing, learning and speaking. Without it we would not know what to do with the information we receive through our senses. At its most basic physiological level, memory is "a physical change in the neuronal structure of the brain" (Coe 1996:69). When information is added to our memory it creates new neuronal pathways and connections.

There are three types of memory:

1. Sensory Memory
2. Short-term Memory (STM)
3. Long-term Memory (LTM)

These three types of memory work together, passing information between them to allow us to carry out cognitive processing.

Sensory Memory

Sensory memory, also known as *sensory registers* (Coe 1996:72) or *sensory buffers* (Dix 1998:27), is the first stage of memory. It is an area of conscious memory which acts as a buffer, temporarily storing information received through the senses. Each of our senses has its own sensory memory (Coe 1996:71), e.g. *iconic memory* for visual stimuli, *echoic memory* for aural stimuli and *haptic memory* for touch (Dix 1998:27). This type of memory acts as a temporary storage area for sensory information before it is passed on for processing. Information stored here is unprocessed, i.e. it remains in its physical form and is not decoded (Downton 1991:22). In effect, this means that the information stored here is extremely detailed and accurate. However, because of the limited capacity of sensory memory, information stored here is the most short-lived and is constantly being overwritten. In general, information is stored in sensory memory for anything between 0.2 seconds (Downton 1991:22) and 0.5 seconds (Dix 1998:27) although echoic memory is more durable and lasts for approximately 2 seconds (Downton *ibid.*).

The existence of iconic memory can be demonstrated easily using the concept of persistence of vision – the principle upon which television and cinema work. By displaying a series of separate images in rapid succession, the eye is "tricked" into seeing a single moving image. Similarly, echoic memory can be illustrated by those instances where we are asked a question and we ask for the question to be repeated only to discover that we actually heard it after all. In a manner of speaking, sensory memory allows us to replay information and gives us a second chance to process information. Sensory memory also serves as a route to short-term memory (STM) for the sensory information we receive (Dix 1998:27; Coe 1996:71). However, due to the brief duration of sensory memory, not all perceptions become proper memories (Raskin 2000:18).

The Low-Capacity Channel

Linking sensory memory to STM is what is termed the "low-capacity channel" (Downton 1991:23). This channel serves as a conduit for information passing from sensory memory to STM. In practice, however, this channel has a low transfer capacity, something which is evident from the difficulty we experience in paying attention to many different sources of information simultaneously. In addition to transmitting information, this channel also converts the information from its raw, physical and unprocessed state into symbolic representations which can be used in STM. Indeed, this is where perception occurs. The limited speed with which this information is converted helps to explain the low capacity of the channel in general. This limitation means that the channel is very prone to overloading.

Short-Term Memory (STM)

A popular way of explaining the concept of STM is to describe it as a "scratchpad" or as RAM in a computer (Dix 1998:28; Hill 1995:19). STM is responsible for storing information that we are currently using. It is where we carry out all of our memory processing, encoding and data retrieval. STM allows us to "do" things with information. We can also filter information here and discard information which is no longer needed.

Card *et al.* (1983:38) argue that STM (or working memory as they call it) is really only an activated subset of information stored in long-term memory (LTM). While it is true that STM obtains some of its input from LTM, e.g. stored knowledge, procedures etc., information passed on from sensory memory also provides STM with its input.

In contrast to information stored in sensory memory, information in STM is stored in the form of symbolic representations or schemes (Coe 1996:72). However, like sensory memory, information is stored in STM temporarily. The information is lost, overwritten or replaced after 20-30 seconds (Downton 1991:24), although with practice information can be retained for several hours (Coe *ibid.*). That information is only stored temporarily is due to the limited capacity of STM. In 1956 Miller posited that the capacity of STM is seven chunks plus or minus 2 chunks. This "7±2" rule is universally accepted as fact (Faulkner 1998:34; Coe 1996:72; Downton 1991:23; Dix 1998:28) and is generally true for most people. This can be illustrated using the following sequence of numbers:

0352765994

The average person may find it difficult to remember each digit in this sequence. However, if we group the digits into smaller sequences as we would with a telephone number, each sequence can be treated as a single chunk:

035-276-5994

So, instead of remembering ten separate pieces of information, by chunking the information we reduce the amount of space required to remember them. An interesting property of chunks is that what actually constitutes a chunk depends on individual people and the content of their LTM (Card *et al.* 1983:36). According to Downton (1991:24) the number of chunks which can be stored is independent of the amount of information each chunk contains. We can, therefore, combine small chunks to form larger chunks and so on. For example, letters (small chunks) form words (larger chunks) which can be combined to form sentences (even larger chunks) and so on (Faulkner 1998:73). With sufficient practice and rehearsal in STM, several sentences can be treated as one single chunk.

Long-Term Memory (LTM)

Long-term memory is the final part of our memory system and it is here that information is stored once passed on from STM. Whereas capacity and retention are key factors when discussing sensory memory and STM, they do not apply to LTM as this type of memory is essentially unlimited in its capacity and information is stored there forever (Faulkner 1998:35; Coe 1996:74; Downton 1991:25; Dix 1998:30).

It is widely held that there is really no such thing as "forgetting" information, i.e. information disappears from memory, rather the information is still stored in LTM but that as the memory grows older, the traces which lead to the information and help us locate it become increasingly faint and less defined. The result is that we simply cannot find the information we want (Faulkner 1998:35). Over time, it becomes more difficult to locate information and this can lead to this information being permanently "forgotten". This condition is exacerbated by the fact that information which is most recently and frequently used is easiest to retrieve from memory (Downton 1991:25).

Of greater interest when discussing LTM, however, is its structure and how information is retrieved from it. LTM can be divided into two types: *declarative memory* and *procedural memory*.

Types of Long-Term Memory

Coe (1996:74) divides LTM into two primary categories - *declarative memory* and *procedural memory*. Declarative memory is described as “memory what” (*ibid*). This is memory of events, facts and images. “Memory how” or procedural memory is memory for motor skills, cognitive skills, reflexes, how to do things etc.

Declarative Memory

Declarative memory consists of a number of different types of memory: episodic, associative, lexical, image memory and semantic.

Episodic memory is our memory of events and facts relating to them (Dix 1998:31; Faulkner 1998:37). This memory allows us to reconstruct certain events and remember facts about them. Coe (1996:75) also mentions a specific type of episodic memory which is like a high-resolution snapshot of a particular event that was particularly surprising, emotional or otherwise significant. This is known as *flashbulb memory* and is attributed to momentous occasions, be they good or bad.

Associative memory is the way we remember information using tags with which we label schemes of knowledge (see also page 35).

Lexical memory is what we use to remember the graphical and phonological features of words. This refers strictly to the physical properties of words – the combination of black and white lines on paper, for example. The **meaning** of these words, however, is stored in *semantic memory*.

The term *image memory* can be quite misleading because it does not only refer to physical pictures we have seen and stored in memory but also to mental images which we construct on the basis of events, pictures, situations, words etc. For instance, we can picture in our minds a childhood birthday party or a beautiful place we may have seen on holiday. We can also manufacture mental images without ever having seen the physical object or place in question. This is part of what we refer to as *imagination* and it is a product of our image memory. For instance, we can picture ourselves sipping frozen margaritas on the deck of a yacht in the Caribbean even

though we may have never been on a boat in our lives, nor visited the Caribbean.

This type of memory is more durable and reliable than any other type of memory (Coe 1996:77). When we store information, either in the form of an image or accompanied by an image, we can recall it more readily than information that does not have image associations. For instance, it is easier to remember where the dipstick is on a car if we associate a visual image of its location rather than a verbal description alone, e.g. a long piece of metal protruding from the oil sump in a car's engine.

Semantic memory is our memory for facts, meanings, concepts, vocabulary etc. Semantic memory is our knowledge about the world – a structured record of facts, knowledge, concepts and skills that we have acquired. Semantic memory is structured to allow orderly and reliable access to the information (Dix 1998:31). One model used to explain how this information is structured is that of the network. Classes are used to relate items together and each item may inherit attributes from superordinate classes (Dix *ibid.*). The classes are all linked together in a network. Semantic networks, however, do not allow us to model the representation of complex objects or events. They merely allow us to model relationships between single items in memory (*ibid.*).

Another structure proposed to counteract this problem is the notion of *frames* and *scripts*. This model organises information into data structures. These structures contain *slots* which contain default, fixed or variable attributes. Scripts are an attempt to model the representation of stereotypical knowledge about given situations (Dix 1998:33). This representation allows us to interpret partial descriptions fully. The frames and scripts are then linked together in networks to present hierarchically structured knowledge.

Another type of knowledge representation is the representation of procedural knowledge. This is our knowledge about how to do something. Procedural knowledge is generally represented in a production system consisting of "*if-then*" condition-action rules (Dix 1998:34). With this model, information that is received in STM is matched against one or more of these rules and the associated *action* is determined by the *then* part of the "*if-then*" rule. For example:

• If the traffic light is green	• Then drive through the junction
• If the traffic light is red	• Then stop and wait until it turns green

Procedural Memory

Procedural memory is acquired using a number of processes: motor skill learning, cognitive and perceptual learning as well as classical conditioning, priming, habituation and sensitisation.

Motor skill learning is the means by which we remember how to do physical activities like blinking, moving our fingers, pushing buttons etc.

Perceptual learning is the process of learning **how** to perceive sensory information each time we encounter it. For instance, the first time we try to ride a bicycle, our senses will tell us that we are not balanced and that we are falling. As a result we may pull our arms up to protect our head as we fall to the ground. With practice, however, this sensory information results in us making slight changes in our body position to correct the loss of balance and continue cycling.

Clinical or Pavlovian conditioning is our memory for a response that is caused by a stimulus and a reinforcer. Drawing its name in part from experiments conducted by Ivan Petrovich Pavlov (Fredholm 2001), this type of memory continues even without the reinforcer.

Priming is the process whereby triggers or cues which activate information from memory are stored. Priming memory is short-lived and available only through the sense that activates it – it cannot be activated or accessed by any other sense. Furthermore, this type of memory does not include the subsidiary information such as when and where the memory occurred.

To put this in context, let us take the following example: if you were asked to think of the colour red and were then asked to think of a particular type of flower and then an emotion, you might think of a rose and anger. The word “red” acts as a trigger which is temporarily stored and which activates other information from memory.

Habituation is the process whereby we become accustomed to sensations repeated over time. If we think back to the example of office air conditioning on page 110, we can see that habituation is the memory that allows us to decrease our attention to the noise of the fan after a certain amount of time.

Sensitisation is the process whereby we acquire sensitivities to specific events, situations or actions. If, for example, you were bitten by a dog as a child, the mere sound of a dog barking may provoke an extreme reaction as an adult.

Retrieving Information from Memory

There are two ways of retrieving information from memory: *recall* and *recognition*. In the case of recall, we reproduce the information from memory whereas with recognition, the information presented informs us that we have seen the information before. Recognition is less complex than recall because the information is provided as a cue (Dix 1998:36).

However, because recall actually reproduces the information and not just the knowledge that we have seen it before, it makes sense to try to assist the recall process. When we try to recall information from LTM, we do not know what the information is or what the cues that aid retrieval are (Card *et al.* 1983:82). However, if we place cues in STM we can assist recall. The problem here is that if we add cues to STM we rapidly fill the STM capacity of 7±2 chunks of information. The result is that while we speed up retrieval, we actually slow down processing in STM. In the case of user guides, the text itself can be used to place cues in STM, but if too many cues are added, the reader's progress through the text and task will be slowed down.

Given the fact that we can recognise information far more easily than we can recall it (Preece 1994:118) it is useful to have users recognise the information they need to perform a task rather than recall it. Of course, there is a trade-off between recall and recognition. Whereas recognised information is easily retrievable whenever the information is present and it does not require learning, recall can be much more efficient because information does not need to be located and assimilated. If, however, information is repeated several times, it will in time become automated in procedural memory, and subject to recall rather than just recognition (Dix 1998:34; Raskin 2000: 18-20).

Cognitive Processing

So far we have examined ways in which we can describe the human cognitive system and we have looked at the components of this system as well as some of their capabilities and limitations. We will now examine how these subsystems interact with each other and function together as a whole to make the human cognitive system what it is – an information processor.

This section looks at how and where we use this system, i.e. what we use our cognitive system to process as well as the actual mechanics involved

in using the system. For our purposes, this discussion will be restricted to how we tackle new information and tasks and how we learn.

Cognitive Conscious and Cognitive Unconscious

In his discussion of human cognitive processes, Raskin (2000:11ff) distinguishes between the *cognitive conscious* and *cognitive unconscious*. This distinction is necessary in order to explain the way in which humans actually go about processing information and, perhaps, to shed light on the limitations and anomalies of how we perform tasks.

Human cognitive unconscious essentially refers to information which we are not consciously using or aware of at a given point in time. We can refer to the cognitive unconscious as those things we are only subconsciously aware of but which are not relevant to what we are currently doing (compare this with the notion of relevance as espoused by Sperber & Wilson 1986 and by Gutt 1991). Conversely, our cognitive conscious refers to information, tasks, etc. that we **are** conscious of, i.e., that we are currently using.

Another way of looking at the difference between cognitive conscious and cognitive unconscious is that when we access and process information we are transferring it from our unconscious to our conscious. This transfer of information can be triggered by a stimulus, such as reading a sentence, or by an act of volition. For the purposes of this study, we can say that cognitive conscious broadly correlates to our everyday notion of attention.

Attention

As we mentioned earlier, our sensory system is under a constant barrage of information and input. We mentioned that in order to function effectively and avoid sensory information "overload" it is essential that we be able to filter and group information in order to extract and absorb what is immediately of relevance to us. But why is this necessary? The notion of avoiding overload is true to a certain extent but the underlying principle is that of attention, or to quote Preece (1994:100), selective attention.

Coe (1996:9) describes attention, or rather attentive processing, as processes that involve cognitive functions such as learning, memory and understanding. *Attentive* processes involve higher cognitive functions. This is in contrast to *preattentive* processes which do not involve cognitive processing

and which are primarily a function of sensory input (*ibid.*). So we can see that attention is similar to, if not the same as, Raskin's concept of cognitive consciousness or information that is currently being processed in STM.

The Locus of Attention

We have a certain degree of control over which information we process in STM. In other words, we can, to a certain extent, control which information is the subject of our attention. For instance, we can be driving home and performing all of the processing necessary in order to do this task and we can then start thinking about what we would like for dinner. In this way we can make unconscious information conscious.

Raskin (2000:17) urges caution with regard to using the word "focus" in relation to attention primarily because *focus* can be used as a verb, and as such it implies some aspect of volition or choice on our part. This, he maintains, can lead to misunderstandings as to the true nature of attention. Instead he uses the expression *locus of attention* to refer to the current object of our attention regardless of how it came to be such. In other words, the locus of attention refers to information that is currently being processed in STM. He justifies this shift in terminology from the widely used "focus" on the basis that while we can deliberately shift our attention to another task, issue or subject (e.g. the cocktail party effect), our attention can be focussed for us by means of other stimuli, e.g. people, events or objects. However, in keeping with common usage in the relevant literature, we will continue to use focus as a verb in the following paragraphs. The preceding caveat regarding volition should, however, be remembered.

This lack of complete control is evident from the following examples: if you are told not to think of ice-cream, the likelihood is that you will think of it. Similarly, if you are thinking about ice-cream, you cannot make yourself stop thinking about it and make the information unconscious unless, of course, you shift your locus of attention to something else.

A key feature of the locus of attention is that we have only one (Raskin 2000:24).We have often heard people say that they can do only one thing at a time, particularly when they are busy. Apart from the obvious physiological constraints preventing us from performing tasks (e.g. we cannot make a cup of tea while at the same time vacuuming the carpet) the reason for this is because in general terms we can focus on and process only one thing at a time. This fact is explained by Card *et al.* (1983:42) who explain that the cognitive system is parallel in terms of recognition but serial in terms of its actions or responses. This means that we can be aware of several

things at the same time but we can only do one deliberate thing at a time. An example of this would be trying to hold an in-depth conversation while listening to the radio. So as a general rule, our attention is focussed on a particular thing (Preece 1994:101). This claim may seem less than credible or quite simply impossible. After all, what about people who can continue a conversation while they are driving? The answer is simple, although perhaps not obvious.

Attention & Selection

When discussing the fact that attention can be either voluntary or involuntary, Preece (1994:101) refers to competing stimuli "grabbing our attention". Herein lies the explanation for the apparent existence of multiple loci of attention. Instead of being able to focus on multiple tasks, our locus of attention switches from one task to another. Raskin (2000:24) also acknowledges this point. When describing how events can trigger conscious attention, he stresses the point that we have not gained an additional locus of attention but rather our locus of attention has been shifted elsewhere (see also Card *et al.* 1983:42). Preece (1994:105) later refers to multitasking which is ostensibly the same as what she calls divided attention or multiple loci of attention. Multitasking is, in fact, "continually switching between different activities rather than performing and completing tasks in a serial manner". Both Preece (*ibid.*) and Raskin (2000:16) acknowledge that our ability to perform tasks is sequential, or serial, rather than truly parallel.

But the question arises here of how we switch our attention from one task to another. After all, if we are focussed on one task, how can we switch to another task if we are capable of consciously processing and responding to only one task or stimulus? This is a wide-ranging and problematic question in cognitive psychology and it would be impractical to discuss every aspect of this issue here. Instead, we will discuss the main theories as to how attention switches from one task or stimulus to another.

But before we embark on this discussion of attention, however, it would serve us well to quickly recap on preceding paragraphs. We know that our senses are constantly receiving information and that this information is stored temporarily in sensory memory or registers. We also know that attention fundamentally refers to the active, conscious processing of information at a given time. This means that just one of the numerous sources of information or stimuli is being processed. Attention is, therefore, the process

by which we select information from the mass of incoming information and process it.

Numerous theories have been formulated over the decades to account for the way our attention changes focus to concentrate on various stimuli. Fundamental to all of these theories is the question of what happens to the "unattended" information (Ellis & Hunt 1993:50), or rather the information we are not paying attention to at any given moment. The main approaches to answering this question are grouped under Bottleneck Theories (Gavin 1998:34) and Capacity Models (Gavin 1998:37; Ellis & Hunt 1993:50-52) below.

Bottleneck Theories

Bottleneck theories fall under the categories of early selection and late selection models and they generally revolve around some variation on the notion of filters. Indeed, the idea of filters is a key element of both early and late selection theories. If we cast our mind back to the idea of the cocktail party effect, we will recall that we can filter out stimuli and focus on one particular stimulus.

Early Selection Filter Models

In early selection filter models, we work on the assumption that only one source of information can be processed (Ellis & Hunt 1993:52). Logically, this means that unattended information is filtered out before cognitive processing takes place, i.e. before the information reaches STM. We can see, therefore, that early selection takes place early on in the information-processing chain of events.

Perhaps the most well known early selection filter model is *Broadbent's Switch* (Ellis & Hunt 1993:52ff). Broadbent (1958) proposed that our attention is determined by a filter and a detector located between sensory memory and STM (Gavin 1998:35). Using the idea of a switch means that we process information from one source or channel only in an "all or nothing" manner (Ellis & Hunt *ibid.*). Essentially, if one stimulus is being processed, all other stimuli are effectively blocked out. But if we are blocking all of the remaining sources of information, how do we remain sufficiently aware of our surroundings to be able to operate the switch and shift our attention? How do we decide how, when and where to focus our attention? According to Broadbent, the unattended information is subjected to a pre-attentive

analysis (i.e. it is analysed before we become aware or conscious of it) which examines the physical nature of the incoming information. From our discussion of sensory memory earlier we should recall that information is stored here in a detailed and unprocessed form. This means that any other form of analysis of the information would be impossible before the information is passed on to STM. Information which is selected on the basis of physical characteristics is then passed along the low capacity channel and into STM for processing. The remaining information is left in sensory memory where it decays and disappears after 0.2-0.5 seconds.

Unfortunately, the notion of an "all or nothing" switch does not explain the cocktail party effect. If we are concentrating on one particular conversation to the exclusion of all other sensory input, how can we detect our name being spoken and change the focus of our attention? The audio input arising from our name being spoken is not processed cognitively and as such, the physical representation of the sound in sensory memory has no meaning for us. This problem is also highlighted by Gavin (1998:36) and Ellis & Hunt (1993:54-56) in their discussions of experiments carried out by Cherry (1953) and Treisman (1960). Treisman discovered during experiments involving dichotic listening and shadowing[1] that subjects were able to report back on some of the content of the unattended information and that the unattended information could even affect the performance of the attended, shadowing task (Ellis & Hunt 1993:55). It is obvious from this that the unattended information is subject to at least some limited form of cognitive processing. This presents obvious problems for the application of the basic switch model.

The Attenuator Model

In light of the problems associated with the switch model highlighted in various experiments, Treisman (1964) developed a more flexible theory for the early selection of information. Rather than using a simple, binary switch, Treisman proposed the use of an attenuator. An attenuator is a type of switch which is used to control the volume or speaker balance on stereo or radio equipment. Instead of a simple on/off function, an attenuator controls signals in stages or gradations to allow more or less signal through.

---

[1] Dichotic listening involves subjects wearing headphones and simultaneously listening to two separate audio inputs or speech recordings, the first through one earpiece and the second through the other earpiece. Shadowing involves subjects repeating verbatim the speech that is heard through one of the earpieces.

To give a practical example of this, if we think of the speaker balance on a stereo we know that if we turn the dial (attenuator) to the left, more sound passes through the left speaker channel and less passes through the right speaker channel. If we turn the dial to the right, the opposite happens. So rather than all of one channel and none of the other channel passing through for processing, we have a situation where virtually all of one channel and **some** of another channel are processed. Like Broadbent's switch, all input is physically analysed and filtered before it reaches STM (Ellis & Hunt 1993:57). Unlike Broadbent's switch, however, all of the information passing through the attenuator can conceivably pass through to semantic processing (Gavin 1998:36). The difference here is that some of the information is "turned down". Gavin (*ibid.*) maintains that such a model does not make "cognitive sense" and does not make effective use of the cognitive system because the level of processing carried out on unattended information is not far from full cognitive processing. Such a situation would undoubtedly use up all of the cognitive system's resources. Ellis & Hunt (1993:58) state that the attenuator model is frequently regarded as being too cumbersome to be practical.

Late Selection Filter Models

In contrast to the early selection models outlined above, late selection models operate after information has undergone partial cognitive processing. Information stored in sensory memory is passed to STM where pattern matching occurs. When the information activates its LTM representation, i.e. a match is found for the pattern, a response is initiated. However, the human cognitive system is only capable of handling one response at a time (Ellis & Hunt 1993:58; Card *et al.* 1983:42). This means that instead of being able to focus our attention on one *input* only, we are, in fact, only able to concentrate on one *output*, namely the item of information activated in LTM.

Late selection functions on the basis of all information being passed in parallel into STM. There are obvious limitations as a result of the low capacity channel linking sensory memory and STM, but for the most part, all information is passed through. Given the fact that STM is limited in its capacity, only some of the information can be stored there. The decision as to which information is stored in STM is based on the relative importance of the information. Information relating to the task at hand is assigned greater importance (Gavin 1998:37). Information which is deemed to be important is then subjected to more rigorous processing. While several items of information may be processed, we can organise and handle responses to one of them only.

Preconscious Processing and Semantic Priming

A crucial characteristic of the late selection model is that all information is processed to some extent. As attention selection takes place after recognition of information and at the point where responses are selected, the unattended information can sometimes be put to good use. Although we can only deal with one response at a time, the fact that we partially process so much information has an effect on the content of STM and which parts of LTM are activated.

Semantic priming is a phenomenon whereby in cases where "the stimulus which preceded the current stimulus was semantically related to the current stimulus, the response to the current stimulus is affected by the preceding stimulus" (Ellis & Hunt 1993:60). This means that stimuli can pave the way for subsequent related stimuli. In effect, semantic priming activates or facilitates access to related items of information. This makes the recognition of subsequent information easier (*ibid.*).

Capacity Models

Capacity models work on the assumption that the human cognitive system has a limited, finite set of resources. When we are aware of more than one stimulus, the available resources are allocated in varying amounts to different stimuli. For instance, a complex stimulus will require more of the available resources with the result that there are fewer resources available for processing other stimuli. Accordingly, attention can thus be defined as "the process of allocating processing resources or capacity to various inputs" (Ellis & Hunt 1993:62). This model is demonstrated in what Ellis & Hunt (1993:63) refer to as the *secondary technique and cognitive effort*. This technique is used to show how much cognitive effort (or processing resources) is needed for a secondary task while a person is already performing a primary task. Experiments conducted on this basis showed that the more complex the primary task, the more difficult it is to perform a second task. Interestingly, the greater the cognitive effort required for a task, the better the memory for that task and related information (*ibid.*).

Gavin (1998:38-39) characterises the model as follows: Competing stimuli produce non-specific interference – the two tasks do not directly interfere with each other but they do compete for the same finite set of resources. The more resources that are needed for one task, the fewer resources are available for the other task. It follows, therefore, that we can theoretically perform two tasks provided the total processing demand does not exceed

the available capacity. When total processing demand exceeds the available capacity, performance on one task deteriorates (cf. Raskin 2000:21). The allocation of resources is flexible and can be adjusted to meet the demands of stimuli.

Performing Simultaneous Tasks

So what happens when we try to perform competing tasks? What happens when our attention is divided? Preece (1994:105) distinguishes between primary and secondary tasks. Primary tasks are those tasks which are most important at that particular time. In order to successfully multitask, Preece maintains that we need to be able to switch rapidly between the various tasks. The task currently being carried out is said to be *foregrounded* while the other tasks are temporarily *suspended*.

In principle at least, this appears to be a more than adequate solution to the problem of multitasking. There are, however, a number of problems associated with multitasking and which detract from the apparent benefits of performing multiple tasks at the same time. The first problem is that people are very prone to distraction (*ibid.*). When switching between tasks, our attention is temporarily not focussed, leaving our attention prone to becoming focussed by other stimuli. Also, there is a tendency to forget where we left off when we return to a task with the result that we can return to a task at a different point from where we left it. Another problem associated with switching tasks is that the more intensely we are concentrating on a task, the more difficult it will be to switch our locus of attention to a new task (Raskin 2000:20).

Raskin (2000:21) also states that when we attempt to perform two tasks simultaneously, our performance on each task degrades. This phenomenon is known as interference and can be explained by the fact that both tasks are essentially competing for the same finite amount of processing capacity. An important point to bear in mind here is that Raskin is referring to two tasks which are not automatic and as such require large amounts of processing capacity. But what are automatic tasks and how do tasks become automatic? To understand this we will need to examine the cognitive processes involved in learning.

Learning

As we saw in preceding sections, long-term memory is a vast and virtually unlimited mechanism for storing information. Just as not all perceptions become memories, not all information is stored in memory; it must first be learned. What is more, we cannot "force" information to be stored in long-term memory. Learning can be defined as a relatively permanent change in behaviour as a result of experience. Novak (1998:19ff) distinguishes between two principal types of learning: *rote learning* and *meaningful learning*. Rote learning is, according to Novak, not true learning and is not ideal because it is an arbitrary, verbatim and non-substantive incorporation of new knowledge into a person's cognitive structure. Rather than learning new information and integrating it with existing concepts and relating it to existing experiences, rote learning involves memorising information in a "word-for-word" fashion. And because the information is not related to existing knowledge, it is very difficult if not impossible to become proficient enough to use the information independently and creatively. All that is remembered is the literal message and not the knowledge or meaning behind it (Novak 1998:20).

Meaningful learning, on the other hand, is a "process in which new information is related to an existing relevant aspect of an individual's knowledge structure" (Novak 1998:51). This type of learning is the best type of learning because it is a non-arbitrary, non-verbatim and substantive incorporation of new knowledge into a person's cognitive structure. This means that information learned is linked with existing knowledge and concepts and with higher order concepts. Essentially, the learner understands the information, how it relates to existing knowledge and ultimately the learner understands how to use the newly acquired information. Meaningful learning, however, requires a conscious decision on the part of the learner in order to take place (*ibid.*). It cannot be forced, merely facilitated.

The Basis of Knowledge

Although it may seem obvious, when we learn we are acquiring new knowledge or, more specifically, meanings. But what is meaning? Novak presents the terms *concept* and *proposition* as the basic units of knowledge and meaning. Concept is defined as "a perceived regularity in events, objects or records of events or objects designated by a label" (Novak 1998:22). A concept is, therefore, a name given to a series of things or events that we

observe to have certain similarities. Propositions, on the other hand, consist of two or more concepts which are combined to form a statement about an event, object or idea. They are the principal units of meaning. Propositions can be valid, invalid or nonsensical. Novak (1998:40) likens propositions to molecules while concepts are compared to atoms.

The Components of Meaningful Learning

There are, naturally, several stages involved in meaningful learning and it would be impractical to discuss all of them here. However, we will discuss three of the key components of learning here: *concept learning*, *representational learning* and *propositional learning*.

Concept Learning

If concepts are the smallest meaningful units of knowledge, it holds that they are the first thing that will need to be acquired in order to learn new information. This process is called concept learning and it is the subject of much debate. There are those such as Piaget who maintain that in order to acquire a new concept, we must first perceive the regularity to which the concept refers before we can acquire the label. This may or may not be true. Theorists such as Vygotsky (1962) believe that having a label already stored in memory can actually assist in the acquisition of the concept.

Representational Learning

Representational learning is a type of meaningful learning where we learn a word, sign or symbol that acts as a label for an object or event. An example of this is the way we learn proper nouns. As mentioned earlier, we can either learn labels before we learn the concepts or we can learn concepts before we learn the labels. Similarly, representational learning may take place either before or after concept learning. However, representational learning on its own is insufficient in terms of meaningful learning because the concept is not acquired and there is no meaning or interrelationship with other knowledge.

Propositional Learning

If concepts are like atoms and propositions are similar to molecules, it follows that out of a small number of concepts it is possible to create a large number of combinations (or molecules). In practical terms, the meaning we acquire for a concept is formed from the composite of all the propositions we know that contain the concept. The more propositions we have that contain the concept in question, the richer our understanding of the concept will be (Novak 1998:40). And so, propositional learning is the process of linking, integrating and associating concepts to provide richer and more detailed meaning. The processes by which we acquire and combine new concepts and propositions are described below.

Acquiring New Information

There are two primary ways in which we acquire new knowledge: *concept formation* and *concept assimilation*. Concept formation involves constructing meanings for words from observed regularities. To illustrate this, imagine we have seen lions, tigers, cats and dogs and they all eat meat. When we learn that this common activity makes them carnivores, we form the concept of *carnivore*.

With concept assimilation, we acquire meanings for concepts by associating them into propositions which contain already known concepts. This can be illustrated using the example of *scone*. We know *scones* are a type of *bread* which in turn is a type of *food*. Here the new concept – scone– is subsumed beneath the concept of bread which is in turn subsumed beneath the concept of food. In this example, food and bread are subsuming concepts. The process of subsumption results in changes not only to the new concept but also to the concepts which subsume it. Consequently, information recalled about scones may be slightly different from that which was originally learned. Similarly, if over the passage of time, the concept of scone is forgotten or cannot be described adequately – a process known as *obliterative subsumption* – it will have modified associated information sufficiently to provide enhanced information about that particular concept area. So while we may not remember the precise details of the information we learned, we will still recall the knowledge it produced as a result of being learned. The process of concept assimilation is never fully finished because we are continually adding or associating new concepts with existing ones (Novak 1998:59-61).

Learning Processes

In general there are two main approaches to learning theory: behaviourist and cognitive. Behaviourist learning theories focus on objective, quantifiable behaviour rather than on mental acts which we cannot observe. They are concerned with the connection between actions, the role of reward in behaviour etc. Cognitive learning theories focus on mental acts such as conceiving, believing, expecting etc.

According to Coe (1996:34) we learn using a combination of behaviourist and cognitive approaches. The components of learning include:

- experience
- schemes
- habits
- reinforcement
- interference

Experience & Schemes

We learn from experience. When we are met with an experience or situation we either create a new scheme or we use/modify an existing one. Thus any information provided, for example, in an instructive text such as a user guide, must either take advantage of readers' existing schemes or help them create new schemes quickly and easily. The easiest way to leverage existing schemes is to give examples based on schemes they already have.

Habits

Habits are learned connections between a stimulus and a response. The strength of the connection is called the habit strength. Related habits are grouped into habit families, each of which has a hierarchical pecking order. The most effective habits which we tend to use first or most frequently are located higher up in the hierarchy.

New habits can be introduced by comparing and contrasting old habits with new habits or building on existing habits. We will discuss habits in more detail below.

Reinforcement

Reinforcement is the process of using events or behaviours to produce learning. These are known as reinforcers and they can be either positive or negative. If we relate this idea to software user guides, one possible example of positive reinforcement would be if it tells a new user about using the keyboard shortcut CTRL-P to print a document. Each time the user does this, the document is printed, thereby reinforcing the knowledge that the shortcut works and is another way of printing a document.

Negative reinforcement involves the removal of an unpleasant or undesirable situation. For instance, if a user accidentally deletes all of the text in a document, pressing CTRL-Z will undo the previous action and restore the deleted text. This removes the undesirable condition and reinforces the user's knowledge of the undo function. In this way, the information in a user guide reinforces learning of functions by allowing users to do something useful or to correct problems.

As well as the positive/negative dichotomy, reinforcement can be divided into the following types:

Continuous

Continuous reinforcement takes place each time an event occurs. This is the quickest way of promoting learning and it establishes very strong expectations for the user that reinforcement will always take place. The result of this is a dependence on the part of the user and consequently if the reinforcement stops, the learning stops too.

Intermittent

Intermittent reinforcement, as its name suggests, is the occasional reinforcement of learning. While initial learning is slower, learning will be more autonomous and will continue even if the reinforcement stops.

Vicarious

This type of reinforcement involves learning from the experiences of others. In other words, we learn to perform those tasks we see others rewarded for. Similarly, we learn not to perform those tasks we see others punished for. A classic example is of a vending machine. If we see someone insert money and obtain two cans instead of one, we will be inclined to do the same. Conversely, if we see a person insert money but not receive anything, we will not risk the same fate ourselves.

However, with vicarious reinforcement, the learning continues until such time as a new observation is made. Returning to the vending machine, if we subsequently see another person successfully use the seemingly broken machine, we will change our knowledge and actions to incorporate this new learning.

With reinforcement, we need to adapt the type and amount of reinforcement according to the audience and medium being used. For example, vicarious reinforcement is not particularly useful for user guides and more advanced users may take exception to frequent and unnecessary reinforcement.

Interference

Frequently, existing habit families will interfere with new learning. Of course, the opposite is also true. Applying this idea to a situation in which a user guide might be used, we can use the example of a user who is proficient in using *Microsoft Word* learning to use *Corel WordPerfect*. The commands and procedures used to create and update a table of contents in *Word* may interfere with those of *WordPerfect* because the user has developed habits from using Word.

On the other hand, interference between existing habits and new learning can sometimes be positive. Returning to the idea of a *Word* user learning to use *WordPerfect*, some of the habits learned from Word will actually aid learning the new program, e.g. creating mail merges, the idea of creating templates or performing spell-checks.

Learning & Cognitive Processes

Behaviourist approaches to learning cannot account for the bulk of human learning because they do not take into account the role of cognitive processes. Cognitive approaches emphasise the cognitive processes of the individual and the social environment in which learning takes place. So, rather than learning being a mechanical series of responses to stimuli, it is "mediated" by cognitive processes. With cognitive processes mediating between stimuli and the individual's behaviour, learning is said to be social (Gavin 1998:119). Furthermore, referring to a study carried out by Tolman (1948), Gavin explains that by a process known as latent learning, we unknowingly learn and assimilate information. We may never act upon this information or knowledge unless we need to or want to. This is where the expectancy

of an outcome or consequences of behaviour determines whether we carry out a task or respond to learning. According to Gavin (1998:120), this expectancy is affected by:

The Locus of Control

This locus can be internal or external. Where the locus is internal, a person believes that he or she can control his or her own destiny or fate. The locus can also be external which means that a person believes that his or her own fate is controlled by external factors. The ease and speed with which we

learn depends on whether the locus of control is internal or external. We are more likely to assimilate a cause and effect if we are "responsible" for that cause.

Learned Helplessness

This is the belief that we cannot alter our fate and as a result, we give up. Gavin (*ibid.*) defines this as the *expectancy of non-escape*.

Explanatory Style

This refers to the way an individual perceives events, especially negative ones. Essentially, a pessimistic explanatory style leads to learned helplessness while an optimistic style empowers individuals and allows them to process alternative responses. We also learn by observing the actions of others. This is the same principle as vicarious reinforcement.

Habits and Automatic Processing

The previous discussions of attention and models of selecting information included various models such as the early selection model, the late selection model (including the attenuator model) and capacity models. Whereas the early selection model effectively precludes the simultaneous execution of multiple tasks and the late selection model places restrictions on the type of tasks that can be performed simultaneously, the capacity model provides sufficient scope to allow for the execution of multiple, fairly high level tasks simultaneously. What is crucial here is not the nature of the task *per se*, because we can often perform two complex tasks at the same time, but rather the way in which we can perform the task. More specifically, capacity theory tells us that the performance of two simultaneous tasks depends on the amount of cognitive resources used by each task. Ultimately, this is a function of how we have learned how to do the task. Essentially, if we have learned a task "well", we need fewer cognitive processes to perform it.

Over time, tasks we have learned well may become automatic, i.e. they are processed automatically.

A good way to begin understanding automatic processing and tasks is to look at habits. It would be virtually impossible for us to function without some form of habit to aid us in our day to day activities. We discovered earlier on that habits are learned connections between a stimulus and a response. Habits are essentially automatic tasks and they are carried out without any conscious effort on our part. Indeed, it would require significant effort for us not to develop habits and, once they have developed, to prevent ourselves from performing these tasks (Raskin 2000:20,21). Of course, tasks do not spontaneously become automatic. They require practice, rehearsal[2] and repetition. Indeed, the more a task is practised and the more familiar we become with it, the easier it becomes and the less we have to concentrate on it (Gavin 1998:33; Raskin 2000:18,19). Take for example, a student learning to play the piano. When starting to play, the student may find it difficult to play without looking at the keys. However, with practice, the student learns the position of the various keys simply by touch. Eventually, the student will be completely comfortable playing without ever looking down at the keys. Similarly, cognitive tasks can become automatic. An automatic process, therefore, occurs without any intention to do it. Such automatic tasks can even override non-automatic tasks. This can be illustrated by means of what is known as the *Stroop Effect* (Faulkner 1998:48). This effect was used to show that the act of reading is so automated, i.e. we have practised it to such an extent, that it takes virtually no conscious effort and can actually take priority over conscious tasks. The experiment involves writing the names of different colours, e.g. red, green, blue, yellow, black etc. but using different colour inks to those described by the word

When subjects were asked to name the colour inks used to write the words, it was found that the information obtained by reading the words overruled or seriously hindered the information which came from recognising the colour of ink. This shows that for the majority of people reading has been practised to such an extent that it has become an involuntary or highly automatic task. In a sense, reading has been "over-learnt" (Gavin 1998:41) and it interferes with other tasks. In practice, however, what this means is that the process of reading requires little in the way of conscious action and few of the resources provided by the cognitive system. So what

[2] Rehearsal is the process of repeating information repeatedly in one's mind.

we have, therefore, is a process whereby we can process and assimilate information without expending excessive amounts of resources. This leaves the remaining resources free to be allocated to other tasks, e.g. learning and solving problems.

Reasoning & Problem-Solving

In the preceding discussion of automaticity and simultaneous tasks, we concentrated on the fundamental cognitive functions involved, namely attention and selection. We saw that with practice, our knowledge about how to perform tasks becomes automated and requires less cognitive effort. However, this discussion omitted a crucial fact: we do not necessarily have the skill and procedural knowledge needed to perform these tasks from the outset.

Referring back to the discussion of procedural knowledge on page 120, we know that procedural knowledge is basically knowledge about how to perform tasks. But what happens when we encounter a task for the very first time? Having never been confronted with it, we do not have a procedure for achieving our goal. This is what is termed in cognitive psychology as a problem. Coe (1996:99) defines a problem as "a goal for which we have no attainment strategy".

As Anderson (2000:240) notes, human cognition always takes place with a view to achieving goals and removing the obstacles that prevent the achievement of goals. So in a sense, virtually all human activity is either problem-solving or originated in problem-solving. Anderson (2000:241) continues to say that there is a tendency to use the term "problem" only for "original difficult episodes" but in reality, all procedural knowledge stems from the resolution of problems. The fact that some tasks become automated does not mean they are not responses to problems.

Reasoning

While problem-solving revolves around finding strategies for dealing with new tasks or experiences, it depends on our ability to reason, i.e. how we make use of existing knowledge to draw conclusions or infer something from either implicit or explicit premises. Reasoning is a cognitive process whereby we reach a conclusion and then determine whether it is valid or

invalid by applying certain logical criteria (Dix 1998:38; Coe 1996:109; Ellis & Hunt 1993:290-291). There are three types of reasoning: *deductive*, *inductive* and *abductive*.

Deductive Reasoning

Deductive reasoning attempts to arrive at logical conclusions on the basis of a set of explicit premises. With this type of reasoning, we seek to derive particular conclusions from general truths which we know or believe to be true (Coe 1996:110).

As a whole, reasoning is based largely on the study of formal logic (Ellis & Hunt 1993:291) and involves reaching conclusions based on general assumptions (*ibid.*). However, there a number of points that must be remembered here. Firstly, humans do not always reason logically (Ellis & Hunt 1993:295; Hill 1995:20). Secondly, a conclusion may be correct from a purely logical point of view, but it may have no relation whatsoever to the real world or how we view it. An example of this would be the following:

- If the sun is shining, it is night time.
- The sun is shining.
- Therefore it is night time.

We can see from this example that a logically correct deduction is not necessarily true in terms of the real world.

Inductive Reasoning

Inductive reasoning is the process of generalising information from specific cases we have seen and then applying this general knowledge to cases we have not seen. For example, we can generalise from our past experiences that all birds have beaks because every bird we have seen had a beak. However, this method is unreliable in the sense that assumptions made using this method cannot be proved to be true – they can only be proved to be false. This is because we cannot possibly see every bird that ever lived or will live. And so there is a risk – theoretically, at least - that the next bird we see may not have a beak. However, each bird we see that *does* have a beak serves as a positive instance which reinforces our generalisation.

To put this in context, let us assume we say all cars have four wheels. Every car we have ever seen has had four wheels. Each one of the many

cars we saw simply served to reinforce this generalisation until one day we see a *Robin Reliant* which has only three wheels. This challenges our belief that all cars have three wheels. However, the fact that we have seen so many examples of four-wheeled cars means that while our generalisation has been proved to be false, we are unlikely to discard it because cars almost always have four wheels. We may modify this generalisation to say that cars have three or four wheels but usually four. This illustrates the iterative nature of reasoning whereby we derive, apply and modify general truths (Coe 1996:110).

In spite of the unreliability of inductive reasoning, it serves as a useful method for maintaining information for general purposes and allowing us to make fairly stable generalisations about the world.

Abductive Reasoning

Abductive reasoning refers to the way we derive explanations from facts (Hill 1995:21; Dix 1998:40). Essentially, abduction involves us trying to find the causes of or explanations for things we see.

Let us suppose that Bob always walks to work when his car is broken. If we were to see Bob walking to work one morning, we might infer that his car is broken. As plausible and possible as this may seem, it is unreliable because Bob may simply have decided to walk to work because he wants to get more exercise.

Despite the fact that abductive reasoning is very unreliable, people frequently infer explanations using this method. Indeed, beliefs acquired using this method will persist until events occur to show that an alternative is true (compare with vicarious reinforcement on page 134).

Theoretical Approaches to Problem-Solving

There have been numerous theoretical approaches to problem-solving over the decades but the most influential approaches are, perhaps, the stimulus-response approach, Gestalt theory and the information processing approach (Ellis & Hunt 1993:287; Gavin 1998:104).

Stimulus-response Approach

This approach assumes that learners approach problems with a number of existing habits of varying strengths arranged into habit-family hierarchies. Based on the principle of habit formation, this approach maintains that certain habits are used to tackle problems at the expense of other habits; the chosen habits are strengthened while the others are weakened. This approach, like that of the Gestaltists is internalised and does not provide enough evidence to prove reliable.

Gestalt Theory

The Gestalt theory rests on the fundamental assumption that the way people solve problems depends on how they perceive and structure their problem environment (Ellis & Hunt 1993:288). Gestaltists maintain that humans' ability to reorganise and recombine their perception of the problem allows problems to be solved. This approach identifies four stages in problem-solving (Ellis & Hunt *ibid.*; Gavin 1998:105): *preparation*, *incubation*, *illumination* and *verification*. Gestalt theorists such as Wertheimer (1959) maintain that problem-solving is both a productive and a reproductive process. Reproductive thinking draws on or reproduces previous experience while productive thinking involves restructuring, insight and the creation of new organisations or *Gestalts* (Dix 1998:44; Gavin *ibid.*).

Preparation	A person gathers information and makes initial attempts to solve the problem.
Incubation	The problem is left for a while and other tasks are carried out.
Illumination	The solution occurs to the person suddenly after incubation.
Verification	The solution is checked to see that it actually works.

Table 2: Stages in Gestalt approach to problem-solving

Gavin (*ibid.*) makes the point, however, that because Gestalt theories are based on introspection, there is insufficient proof that all of these stages occur in all cases. Consequently, such theories "lack the comprehensiveness necessary for a good theory" (Ellis & Hunt *ibid.*). While Gestalt theory ultimately proved unsatisfactory in terms of explaining problem-solving, it

did prove useful because it marked a shift away from previous stimulus-response approaches (Ellis & Hunt *ibid.*) and towards the information-processing theory which is now so prevalent.

Information-Processing Approach

In contrast to both the Gestalt and stimulus-response approaches to problem-solving, the information-processing approach attempts to model problem-solving from a computer perspective. The aim is to formulate steps and rules which are involved in solving problems in order to produce an abstract model of the process (Ellis and Hunt 1993:289). Developed by Newell & Simon (1972), the information processing approach places problems in what is known as a *problem space*. In his book written with Card and Moran, Newell defines a problem space as

> …a set of states of knowledge, operators for changing one state into another, constraints on applying operators and control knowledge for deciding which operator to apply next. (Card *et al.* 1983:361)

We can further explain the problem space concept by saying that it consists of various states of a problem (Anderson 2000:242; Dix 1998:41). A state in this regard is a particular representation of the problem. In solving problems, a person starts out from what is called the *initial state* where the problem is unsolved and navigates through the problem space until the *goal state*, where the problem is solved, is reached (Anderson *ibid.*; Gavin 1998:106; Dix 1998:41-42).

In moving from the initial state, the person changes one state into another using problem-solving operators. Operators are possible moves which can be made in order to change one state into another or to divide goals into sub-goals. Basically, problem-solving involves finding a series of operators which lead from the initial state to the goal state. One crucial feature of the problem space model is that it takes place within the cognitive processing system and as such, is limited by the capacity of STM and the speed with which information can be retrieved from LTM. It is also important to note that there are different problem spaces for different tasks and that problem spaces may change over time as a person becomes more familiar with the task (Card *et al.* 1983:87). As states are converted into other states, it may be possible to use any one of a number of possible operators. The

challenge here is to select the appropriate one to form one of a series of stages which make up problem-solving.

Stages in Problem-Solving

In the information-processing approach, a problem is placed in a problem space. We know that the problem space consists of the various states of the problem, namely the initial and goal states as well as the states in between. But how is it that we arrive at the goal state? To understand this, we need to examine the various stages in problem-solving.

1. Identifying and Understanding the Problem
2. Devising and Selecting a Strategy
3. Carrying out the strategy
4. Checking whether the strategy actually worked.

Much of the difficulty encountered in problem-solving stems from a failure to fully comprehend the problem and to recognise its features. If we do not fully understand a problem, we cannot hope to solve it (Ellis & Hunt 1993:266). Consequently, the way we perceive or "frame" the problem ultimately plays a decisive role in our success (Coe 1996:48). There are four principal types of knowledge which affect the way we interpret a problem.

Factual knowledge	Consists of rules, categories and representations of the problem
Semantic knowledge	Conceptual understanding of the problem
Schematic knowledge	Provides an infrastructural understanding of the problem; the various related issues and factors and how they relate to each other
Strategic knowledge	An understanding of how to build strategies for solving problems within the overall problem area

Table 3: Types of knowledge affecting the understanding of problems

Once we have identified and understood the problem, the next stage is to formulate and select a strategy for solving the problem. In this stage, we attempt to find a possible solution to a problem. Indeed, we may even

formulate several possible solutions to a problem (Ellis & Hunt 1993:267). Again, we have several strategies for devising solutions.

Random Trial-and-Error

This process involves the random selection and implementation of operators until the goal state is achieved. Naturally, such an approach can be extremely inefficient and time-consuming

Hill Climbing

This involves gradual movements away from the initial state and towards the goal state. This approach can result in misleading choices and increase the time needed to reach the goal state.

Hypothesis Testing

Hypothesis testing is similar to trial-and-error but instead it is a purely cognitive activity. Rather than testing a large number of random possible solutions, a smaller, more considered range of solutions is tested. Knowledge of the problem area is used to restrict the selection of possible solutions to ones which have a realistic chance of success.

Algorithms

These are sets of precise rules or procedures which guarantee the successful solution of the problem. Algorithms take a set of rules and processes that have worked in the past and apply them to the current problem.

Heuristics

Heuristics are similar to algorithms but differ in the fact that they do not guarantee success. Rather, they are general guidelines or "rules of thumb" (Ellis & Hunt 1993:267; Coe 1996:123). They are loose sets of procedures which can assist in solving a problem. One useful heuristic is the means-end analysis. This is a complex process whereby a person works on a single goal at a time. If it is not possible to reach the goal state at present, the goal is divided into sub-goals with the aim of removing the obstacles to the achievement of the primary goal (Anderson 2000:253; Gavin 1998:103).

Once we have selected and implemented a solution, the next stage is to see whether it worked. Here too, a clear understanding of the problem and also the solution is needed in order to determine the success of the solution. This essentially involves some form of judgement as to the effectiveness of the solution.

Difficulties Involved in Problem-Solving

Despite the wide range of strategies and processes involved in problem-solving, the seemingly straight-forward task of solving a problem is not without its difficulties. One of these difficulties is known as a *problem-solving set*. This refers to the fact that we have a tendency to view problems through our own experience which prevents us from looking at problems in novel or inventive ways. *Confirmation bias* refers to our tendency to search for solutions that confirm our view of existing ideas. We are reluctant to accept ideas which are different to or which challenge our beliefs. *Functional fixedness* refers to our inability to see the flexibility of objects' functions. Finally, if we think back to our previous discussion of habits we can see that our past experiences can play a part in how we solve problems. *Negative transfer* is the phenomenon whereby existing habits and experiences inhibit or prevent us from solving problems.

The Transition from Problem-Solving to Skill

While we are learning, we are constantly solving problems. When reading a user guide, for example, we perceive new information and we process it in order to understand not only the instructions in the text but how they relate to the task at hand. Indeed, all new tasks start off as problems which we need to solve (Anderson 2000:240-1). As we become more proficient and knowledgeable, the extent of problem-solving diminishes as we learn how to perform tasks and these tasks become automatic. As tasks become increasingly automated, they require fewer cognitive resources and can be performed more quickly (Raskin 2000:20-21).

Preece (1994:164-5) maintains that the transition from problem-solving to skill involves converting declarative knowledge into procedural knowledge or skills. However, in order to perform this conversion, users need sufficient declarative knowledge before the associations can be made. Only when this repository of declarative information is present and the associations established between items of information, can a task become automated or procedural. The course of this transition from problem-solving to skill can be mapped by means of user curves.

User Curves

User curves are a way of visualising the progress of learners and helping to determine their needs and achievements. Coe (1996:44ff) compares a user's approach to learning how to use new technology with Erikson's stages of development (Erikson 1965:239-266). The *Technical Information User Curve* is introduced to help plot the stages in learning technical information from entry level users to power level users.

Entry-level users perform only what they are instructed to and achieve the direct results of those instructions. At this early stage, users can only comply with direct instructions and they are only concerned with their immediate circumstances.

Beginners focus on the basic aspects of the task at hand and begin to build simple symbolic models. They cannot yet apply these models elsewhere.

Intermediate users learn more about the basics and develop an understanding of the logic and principles involved.

Power users have a deep understanding of the product and can apply this knowledge to new tasks.

Another unique type of user is the *outsider* who does not participate in the learning process. Outsiders could be occasional users who do not want or need to learn how to use the product, preferring instead, on the rare occasions that they need to use the product, to learn a single task well enough to perform the immediate task. This *ad hoc* use and learning cannot be charted on the curve.

Users can approach this learning curve at any stage and, with practice, they can move up through the various stages. They also can choose not to participate in the curve and remain at the outsider stage.

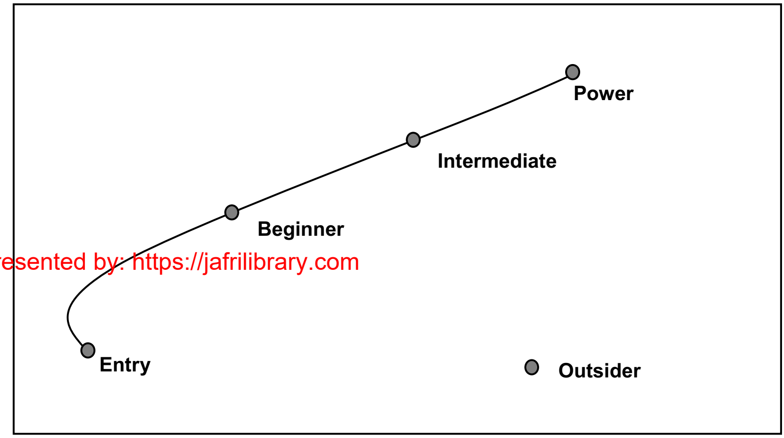

Figure 6: Technical User Information Curve

Although Erikson's Psychosocial Curve aids visualising a person's social development, it can also be applied to user learning in conjunction with the *Technical Information User Curve*. The stages on this curve are:

Trust

Users need to trust in the information and be confident that it will work. Otherwise users are likely to stop there and then (cf. Schneiderman 1998:12-13). At this stage, users follow instructions to the letter. For users to proceed beyond this stage, they need to have faith in the instructions and be confident that they will have the anticipated outcome. Users need constant reassurance and need to feel that their environment is predictable and reliable.

Self-confidence

Users begin to feel less exposed and more confident as regards their abilities. They come to have more faith in the information contained in a user guide. This is a critical stage in the development process because if at any stage the trust built up is broken, i.e. users are given incorrect or faulty information, the user will lose confidence in the instructions and the product and may not use either again.

Initiative

At this stage users are confident and knowledgeable and can comfortably use the product without assistance, prompting or recourse to the user

guide. The user guide becomes a useful but non-essential source of information which users can refer to if necessary.

Competence

With more experience, users become more confident and begin using knowledge for different tasks and applying skills in different contexts.

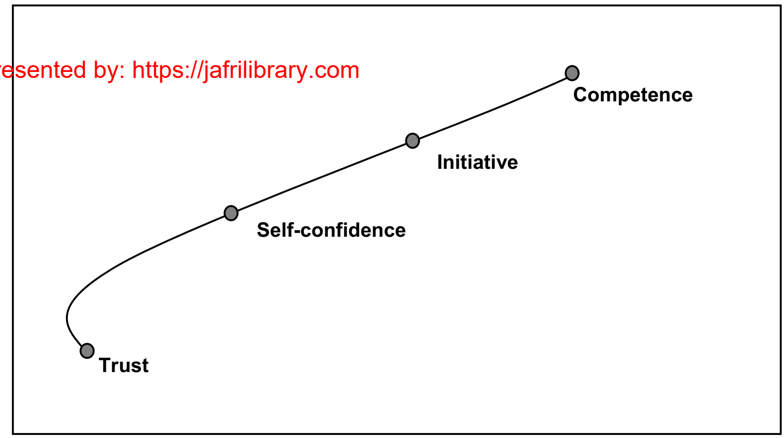

Figure 7: Erikson's Psychosocial Curve

If we compare the two curves we can see the psychological states and information requirements of users as they progress from novice to expert. The aim of a user guide is, therefore, to guide users along these curves through the various stages of learning. In order to do this, the user guide must take into account the various cognitive factors which affect how users process and store knowledge.

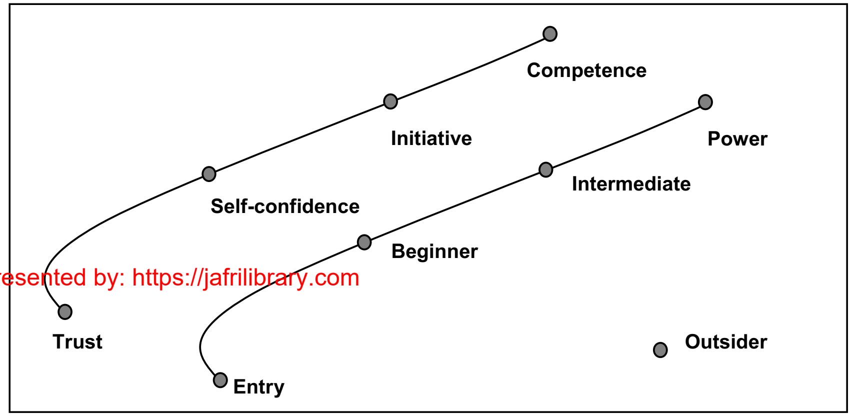

Figure 8: Comparison of Erikson's Curve and the Technical Information Curve

Conclusions

This chapter began by defining usability as the *ease* with which users can perform tasks *effectively* and *efficiently* and their level of *satisfaction* with their work. Having discussed the various factors which contribute to usability, we set about understanding the nature of usability. Usability means that products must take users' mental abilities and limitations into account. To understand usability, it was necessary to explore the various systems and processes that make up human cognition. Human cognition is likened to an information processing system in that humans receive data or information from their surroundings which they then process. These processes enable humans to make sense of the information they require and to decide what, if anything, can be done with it. If a response is necessary, our cognitive processes help us select the appropriate course of action. Sight, the human sense which provides the most important sensory input in relation to user guides, was discussed and its capabilities and limitations were outlined.

We have seen that sensory input is subjected to several, often complex processes in short-term memory (STM). It is here that we decide on courses of action and from here that information passes on to long-term memory (LTM). Armed with this knowledge, we can see how it is that printed words are converted into information that we can use and how this information needs to be processed in order for us to remember or learn it. Of course, the flow of information from STM to LTM is not automatic,

guaranteed or even efficient. The chapter outlined the broad phases involved in turning ephemeral information stored in STM into lasting knowledge stored in LTM through learning. Some of the various obstacles to this process were also presented.

This chapter provided us with the foundations upon which we can present a more detailed discussion of usability and how we can take human cognitive abilities and limitations into account to ensure that when we translate user guides, they are as usable as possible. The task facing us now is to take this knowledge and put it to practical use in the translation of user guides. The following chapter will look at how we can engineer texts to take advantage of the strengths of human cognition and compensate for its limitations.

Chapter 4
Usability Engineering

> Accommodating the diverse human perceptual, cognitive, and motor abilities is a challenge to every designer…the presence of a computer is only incidental to the design; human needs and abilities are the guiding forces. (Schneiderman 1998: 18)

This chapter deals with usability engineering and how we can adapt user guides or "interfaces" to suit humans while making their work as simple and undemanding as possible. In essence, usability engineering allows us to put our understanding of human cognition into practice. The chapter will discuss how an understanding of cognition can inform the way we look at the translation of user guides to make them more usable and to ensure that users can work with a product effectively and efficiently. First of all it is necessary to define what we mean by interface and then apply this definition to texts. By treating user guides as a form of interface between users and a software application, we are able to draw on various principles of interface design and set clear usability goals which will guide the translation process. We will then explore ways of examining user guides from an interaction point of view to identify areas where we can make improvements.

This chapter will then introduce Iconic Linkage (IL) as one possible method for improving the usability of user guides. The chapter will present the origins and nature of IL and proceed to discuss the potential benefits of IL and how it can improve usability.

Interfaces

When we speak about *user interfaces* many people assume we are referring specifically to the graphical user interfaces (GUI) of modern computers. While GUIs are perhaps one of the most common and recognisable types of interface, they are precisely that – types of interface. The reality is that

not all interfaces have windows, icons and menus: interfaces can be found on VCRs, mobile phones, digital watches, ATM machines and even microwave ovens. It is very easy to give examples of interfaces but actually defining interfaces is another matter. Card *et al.* (1983:4) state that it is easy to locate the interface between computer and human simply by starting at the CPU and tracing "a data path outward… until you stumble across a human being". This, however, by the authors' own admission is less than clear and we are left with no real clue as to the nature or boundaries of the interface.

Faulkner (1998:54) maintains that the human-computer interface mediates between the user and the computer system. Again, this is somewhat vague. Perhaps we should look to the function of the interface in order to understand what an interface is. Bødker (1991:77) proposes that "the basic role of the user interface is to support the user in acting on an object or with a subject". She continues by saying that a good user interface allows users to focus on the task at hand rather than on other objects or subjects. So, like the software system itself, the purpose of interfaces is to allow us to do something – in this case, to use the system. In other words, an interface is a tool or a means to an end. Such a view is echoed by Raskin (2000:2) who defines interfaces as "the way that you accomplish tasks with a product". Perhaps one of the clearest and most useful definitions of an interface is that provided by Preece (1994:13):

> The user interface of a computer system is the medium through which a user communicates with the computer. […] the user interface can be thought of as those aspects of the system with which the user comes into contact both physically and cognitively (Preece 1994:13).

Admittedly, the above definition of interfaces is still rather vague in terms of concrete physical details but it is sufficiently detailed in terms of function to allow for variations in the physical characteristics of interfaces and their areas of use (as mentioned above). This flexibility is essential when we consider the case of software user guides. Ostensibly, the purpose of such guides is to teach users how to use a software product. Without such training, many users would not be able to use the software; although a few may try to use it by a process of trial and error although they are likely to have to be less efficient in their learning than those who use the user guide. In other words, user guides facilitate the use of software products and in a sense become part of the human-computer interface. If we were to be very specific, the user guide would be an interface between the user and the software's interface but it is more convenient to simply regard it as an extension of the overall human-computer interface.

Cognetics

Ergonomics is the design of machines to take into account the physical variability of humans. For example, we know that a human cannot possibly be expected to press two buttons positioned three metres apart at the same time. With our knowledge of the human body and the standard level of variation among different humans, we engineer our physical world to suit our needs and capabilities. Similarly, we need to engineer our world to conform to our mental capabilities and limitations. Essentially, what we are talking about is an ergonomics of the mind. This is known as *cognitive engineering* or *cognetics*. In reality, cognetics is a branch of ergonomics but the term ergonomics is used primarily to refer to the physical aspects of human-orientated design.

A key factor which is frequently overlooked by software designers, engineers and even users is that computers are supposed to be tools which assist humans in doing something else. Computers should, therefore, reflect the needs, capabilities and limitations of the humans who use them. As Faulkner (1998:2) says, a computer "has to play up to [users'] strengths and compensate for their weaknesses". Raskin (2000:10) maintains "you do not expect a typical user to be able to multiply two 30-digit numbers in five seconds and you would not design an interface that requires such an ability". But this is an obvious example. Other factors are more subtle and relate to the way we perceive and process information, solve problems, learn and access knowledge – even how we read.

The main challenge facing software manufacturers is to produce systems which people really want, need and can use despite the complexity of the task being performed (Faulkner 1998:129). While decisions as regards what people want and need from a product are usually based on economic factors and made by people other than the actual system designers, ensuring the usability of systems remains the primary focus of cognetics.

It is clear that efforts to ensure usable and successful documentation need to be made from the very start and it would be extremely helpful if the obstacles to usable documentation are overcome from the outset. With this in mind, it would be useful to discuss the goals which must be achieved in order for an interface, in this case a user guide, to be regarded as usable. We will then examine the way in which readers use user guides in order to understand the cognitive processes and effort involved. This will help us identify areas where we can focus our efforts.

Usability Objectives

Faulkner (1998:130-131) maintains that to ensure that a system is as usable as possible, there are three fundamental goals which need to be achieved: *learnability*, *throughput* and *user satisfaction*.

Learnability

Learnability refers to the time required to learn the system or to reach a specific skill or performance level. This objective can be quantified by examining the frequency of errors, the type of errors made etc. Dix (1998:162) expands this category by including the sub-headings of *predictability*, *familiarity*, *generalisability* and *consistency*. Familiarity refers to the way information presented relates to users' existing knowledge which they bring with them to the learning process. Generalisability relates to the ability of users to use the information learned in other situations.

Throughput

Throughput refers to the ease and efficiency of use after an initial learning period. This is quantified by examining the time needed by users to perform tasks, their success rates when performing tasks, the time spent looking for help etc.

User satisfaction

User satisfaction is a subjective goal but it can give an overall picture of how well the system performs in the eyes of users. This can be quantified by asking users to fill out a questionnaire rating aspects of the systems performance etc., for example on a scale from (1) very bad to (5) very good. Schneiderman (1998:15) adds an additional goal which he terms *retention over time*. This is particularly relevant to user guides in that their purpose is to teach users and facilitate their use of the system. Retention relates to how well users maintain their knowledge over time as well as the time needed for learning and frequency of use.

Quesenbery (2001) provides a similar taxonomy of usability characteristics which she terms the "*5 Es*". For a system to be regarded as usable it must be:

1. effective
2. efficient
3. engaging
4. error tolerant
5. easy to learn

Effective

Effectiveness refers to the ability of users to perform tasks completely and accurately. If a user's goals are met successfully and if the user's work is correct, the system is regarded as usable. Quesenbery (*ibid.*) states that effectiveness can sometimes be confused with efficiency. She clarifies this by saying that where *effectiveness* relates to *how well* a task is carried out, *efficiency* deals with *how quickly* a task is carried out. An effective user guide should, therefore, provide correct information which will allow the user to complete tasks successfully.

Efficient

Efficiency is the speed at which users can perform their tasks accurately. In the ISO 9241 standard, efficiency is defined as the total amount of resources expended in performing a task. Quesenbery makes the point that in order to fully appreciate efficiency, it is essential that the notion of a task be approached from the point of view of a user. Whereas a system designer might treat an interaction between a user and a system as a series of small tasks, these tasks are grouped together by users to form one large task. Thus, the procedure for connecting a new printer to a computer might consist of small, efficient tasks such as connecting the cable, installing the drivers, calibrating the printer etc., but if the total time needed to complete all these tasks is greater than that the amount of time a user is prepared to spend, the overall performance of the system is inefficient. In the case of user guides, the information should be clear, digestible and concise. If users have to grapple with unwieldy constructions or ambiguous instructions, the efficiency of the system will undoubtedly suffer.

Engaging

An interface or system is engaging if it is pleasant and satisfying to use. An engaging system will hold a user's attention and make using it a rewarding experience. This characteristic can be affected by such things as the visual design of the interface where the readability of the text as well as the type of interaction can change a user's relationship with the interface and system. The way in which information is chunked also plays a role in how engaging an interface is and helps maximise the resources of a user's short-term memory.

Error Tolerant

In an ideal world, every system and interface would be free from errors and the people who use them would not make any errors. However, it would be naïve to expect users not make at least *some* mistakes. Consequently, a usable system should pre-empt the types of errors a user is likely to make

and either make it very difficult to make these errors or at least provide ways of identifying and rectifying errors when they occur. In the case of user guides, clear and precise information and instructions are essential. Similarly, warnings and precautions should be given in good time to prevent users "jumping the gun" or performing potentially invalid actions.

Easy to Learn

Like the idea of learnability described earlier, ensuring that a system is easy to learn should allow users to quickly become familiar and confident with

the fundamental workings of the system and provide them with a basis for continued learning. A common barrier for users is the typically steep learning curve required for new users. Making systems easy to learn, Quesenbery (*ibid.*) maintains, involves more than just presenting information in a way that is easy to understand. Ideally, users should be allowed to build on their prior knowledge which they bring with them to the interaction (see *Constructivist Approaches* in Chapter 3).

Similarly, if we make the interface as consistent as possible, it should be possible for users to re-use interaction patterns they acquire. This serves to reinforce the information they have already learned as well as that which they will learn. Consistency and predictability are also key factors in Quesenbery's description of the ease with which users learn to use systems. If we use familiar terminology and functionality, users develop expectations as to how the system will behave and this in turn inspires confidence in the user.

We can see that usability is much more than just making sure that everything works with as few problems as possible. It requires a deep understanding of users, what they want to do and what they are capable of doing. Similarly, making sure a product is usable is much more than "tightening bolts, smoothing off rough edges and applying a coat of paint". Rather, it is a much more involved and complex process. According to Dumas & Redish (1993:5) "usability means focusing on users" and in order to develop a product that is usable, it is essential to know and understand people who represent real users.

So how do we set about making systems usable? More specifically, how do we improve the usability of user guides? We need some way of engineering them to make sure that they take into account the problems faced (and posed) by users, particularly their cognitive capabilities. The following sections will examine the process of reading a user guide to learn how to use a software application. This will allow us to identify those aspects of reading user guides which can present problems for readers and it will help us explore ways of alleviating these problems.

Understanding How We Read User Guides

Before we can even contemplate improving the quality of user guides, it is essential to understand why and how users read them. As we discussed earlier, the primary purpose of user guides is to educate users. However, users may have different learning strategies when it comes to software which affect how they read the user guide: some may perform tasks as they read the user guide while others may read the user guide completely before starting to use the software. Others may quickly browse the user guide to find key information before learning the rest as they use the software. These different circumstances are known as *scenarios of use* or *use cases* and they describe how users carry out their task in a specific context.

There are several potential use cases for user guides. Coe (1996:138-140) provides a useful categorisation of these as part of an overall discussion of reading strategies. She identifies two primary types of reading strategy: *reading to learn* and *reading to perform actions*, corresponding to users' desire to acquire declarative knowledge or procedural knowledge respectively. Users may also access a text in a number of ways: they may read a text sequentially or randomly. In addition to the general purpose of reading expressed by the type of strategy and the methods of accessing the text, Coe (*ibid.*) maintains that there are five goals of reading. These are closely related to the type of reading and they describe the method of reading chosen by the user in order to achieve the desired outcome.

Skimming involves reading for the "gist" of a text. This is a declarative goal and can be either sequential or random in nature.

Scanning is another declarative method intended to find specific information.

Searching is a declarative method which is similar to scanning but in this case the reader's attention is focussed on the meaning of specific information.

Receptive refers to reading to fully understand information in a text. It can be either declarative or procedural and it can be either sequential or random.

Critical refers to reading for evaluation purposes. A sequential method that is either declarative or procedural.

Assumptions about Reading User Guides

For our purpose, we will presume that the use case involves users reading the text to perform some task, i.e. they are looking for procedural information. Novice users are, after all, more interested in how to do something (Redish 1988) rather than acquiring a deep theoretical understanding of the underlying principles. We will also assume that users will access information sequentially, at least within a section. Thus, while users may not read every section in sequence, they will at least proceed through each section in the intended order. This is because some readers may, for whatever reason, skip certain sections usually because they may already know some of the information presented or because the information in a section is not immediately relevant to them. Finally, we will assume that the reading goal is receptive and that users want to understand the tasks they are performing.

Having placed the user guide and the user in a specific use case, we need to examine in greater detail the task of reading. Fortunately, a lot of information is already known about the processes involved in reading so we are not forced to start entirely from scratch. Coe (1996:134-135) provides a useful and reasonably detailed summary of the cognitive processes of reading which is based on the following headings:

1. Perceive visual data
2. Recognise words and letters or learn new words
3. Understand the relationship of words to the entire passage
4. Relate the information to a body of knowledge
5. Encode the information
6. Retrieve the information
7. Communicate the information

In the particular use case we have specified for this task analysis, only Stages 1 to 6 are directly applicable because we are using the information immediately in order to perform a task. We are not communicating this information to anyone else although Coe maintains that Stage 7 *can* involve communicating the information to oneself.

Concentrating, instead, on Stages 1 to 6, we can say that in **Stage 1**, perceiving information requires the physiological detection of physical stimuli. This information is then subjected to pre-attentive processing to group the physical marks on the page into shapes which may have a meaning for us. In **Stage 2**, we take these shapes and group them into letters and words. Using a combination of pattern matching techniques (prototype matching, template matching, distinctive features) we match these words with lexical information in long-term memory (LTM). This process identifies the

shapes as words the meanings of which we may or may not have in semantic memory. Where we do not have semantic information associated with the words, procedural memory is activated to provide us with a way of finding out what the unrecognised or new words mean. Such procedures might include how to use a dictionary etc. Once we have located the meaning of the word, we store it in LTM for later use or in short-term memory (STM) for immediate use.

Once we have recognised and identified all of the words, **Stage 3** involves relating these words to each other and with the rest of the sentence, paragraph or text. This requires the retrieval of semantic information for each word and the reconciliation of the various meanings for each one within the context of the meanings of other words. We then chunk this information and combine each chunk with additional chunks. Once we understand the text or passage, we then relate the information it contains to what we already know about the subject. In **Stage 4**, we may create new schemes or modify existing ones. This information is then incorporated into our existing knowledge base which is stored in LTM.

In **Stage 5**, we encode the information in order to integrate it into new or modified schemes. This takes place in STM and the information is encoded as either procedural or declarative information. Ultimately, it is stored in LTM with varying degrees of success.

In **Stage 6**, we retrieve information from LTM. In our specified use case, this would be necessary if, at a certain point in reading the user guide, there was insufficient information available to perform a task. Consequently, the information is either maintained briefly in STM or placed in LTM until the remaining information which is needed becomes available.

These stages govern the cognitive processes of reading in general. There are, however, certain additional factors which distinguish reading in general from reading a user guide to do something. To understand this, consider the following. Once we have performed the processes mentioned above, we are in a position where we have information either in STM or in LTM. The question now arises as to how we are going to use this information. If we are reading to perform a task, our first problem is to decide whether we have sufficient information available to perform the task. As mentioned in Stage 6 above, we may not have all of the information necessary, in which case the information is stored in STM or LTM and we continue reading until we find the required information. Of course there may be cases where readers may not realise that they do not have all of the information necessary to perform a task but this represents a digression from the procedure outlined above.

If we do have all of the information needed, we have to locate the information in LTM. Providing this is completed successfully, we then decide how to perform the task. The question arises as to whether the task can be performed (a) at once, or (b) do we need to divide it into subtasks. If the answer is (a), we perform the task, ensure that we have been successful and continue reading. If the answer is (b), we divide the task into subtasks, perform each one (if possible) and in sequence and then continue reading.

Implementing Usability Measures

In the preceding sections we looked at how humans perceive, process and store information. We also looked at the process of reading and examined those aspects of reading a user guide which place the greatest burden on our cognitive abilities. However, this knowledge is of little value unless we have some way of connecting it with the practical aspects of producing user guides, i.e. writing, translation and editing. The challenge here is to formulate and implement a series of *principles*, *guidelines* and *rules* (Dumas & Redish 1999:53-61; Schneiderman 1998:52ff).

Principles, Guidelines and Rules of Usability

> An interface is humane if it is responsive to human needs and considerate of human frailties (Raskin 2000:6)

In order to ensure that an interface is both "humane" and usable, we need to take the various characteristics of the human cognitive system - as described in the preceding chapters - into account when implementing an interface design. Using principles, guidelines and rules provides a way of selecting key cognitive issues which are of relevance to cognetics and transforming them into practical and workable methods for improving interactions and ultimately the usability of the interface.

Dumas & Redish (1999:52-53) assert that many of the usability problems encountered in practice are dealt with in the wealth of information obtained through interface design research and usability testing. They make the point that interface and documentation design both draw on the same body of knowledge albeit from slightly different angles. They claim that interface design is concerned with designing software to ease the interactions

with users while documentation designers design documentation that "works for users" rather than documentation that just describes the system (*ibid.*). Experts from both disciplines ask similar questions such as "*How do I make sure users' abilities are catered for?*" etc. The mass of knowledge is "distilled into general principles and guidelines" (Dumas & Redish 1999:53; see also Preece 1996:488-491).

Defining a principle as "a very broad statement" that is usually based on research into how people learn and perform tasks, the authors provide an example which applies both to computers and to documentation (Dumas & Redish *ibid.*): be consistent in your choice of words, formats, graphics and procedures. This principle is based on the fact that we learn faster when what we see and do is consistent.

Examples of Principles

Before we begin discussing concrete strategies for improving the usability of our interface, i.e. software user guides, we should first identify the predominant principles of interface design and usability derived from studies on technical communication, cognitive psychology and interface design as described in previous chapters.

In what he terms the "*Eight Golden Rules of Interface Design*", Schneiderman (1998:74-75) sets out a series of principles which play an important role in designing usable, user-friendly and effective interfaces. While there are numerous aspects of interface design that can be drawn upon in interface design, these rules serve as a concise and general overview of the areas that need attention:

- Strive for consistency, use similar prompts, commands, fonts, layout, situations, instructions etc.
- Enable frequent users to use shortcuts
- Offer informative feedback
- Organise sequences of actions so that they have a start, middle and end.
- Offer error prevention and simple error handling
- Permit easy reversal of actions
- Support the internal locus of control, this allows users to feel in charge of the computer and not vice versa.
- Reduce short-term memory load

It is apparent that the principles set out by leading figures in interface design such as Schneiderman share more than a few similarities with those produced for documentation design. Dumas & Redish (1999:61) explain that this similarity is due to the fact that the principles for both are based on creating products that are usable.

Due to the ubiquity of evaluation throughout the development and production process numerous evaluation criteria can also be used as design principles. One such set of criteria was developed by Nielsen (Molich & Nielsen 1990). His *heuristic* evaluation method was developed as a cheap evaluation tool for companies who "couldn't afford or hadn't the resources for empirical usability testing" (Hill 1995:119). Heuristics are general design principles that are usually, but not always, effective (Landauer 1995:283). The following paraphrased list of Nielsen's heuristics represents what are widely acknowledged to be best practice for ensuring usability.

- Use simple and natural language.
- Say only what is necessary.
- Present the information in a logical way.
- Speak the users' language - use familiar words and concepts.
- Minimise users' memory load.
- Be consistent.
- Provide feedback and tell users what is happening.
- Provide clearly marked exits to allow users escape from unintended or unwanted situations.
- Provide shortcuts for frequent actions and users.
- Provide clear, specific error messages.
- Where possible, prevent errors by limiting the number of available options or choices.
- Provide clear, complete help, instructions and documentation.

These principles are widely cited by other sources such as Landauer (1995:283), Hill (1995:119-120) and Dumas & Redish (1999:65).

Examples of Guidelines

Where principles are goals or ideals, they do not say how to achieve the goal. Guidelines are more specific goals and they explain how a principle can be implemented. Dumas & Redish (*ibid.*) state that any one principle can give rise to numerous guidelines although not all of them may be applicable at the same time. Thus, guidelines are derived from principles for a specific context and set of circumstances. Crucially, Dumas & Redish claim that interface design principles and guidelines are only intended to supplement usability testing and that there is no guarantee of a completely usable design, even if all of the relevant principles and guidelines are followed. However, adhering to guidelines makes the incidence of serious usability flaws less likely. Guidelines based on the aforementioned principles might include:

- Always phrase instructions consistently
- Avoid excessively long sentences
- Only use approved terminology
- Use the same formulations and constructions for sentences
- Avoid confusing verb tenses

Rules

Although guidelines are more explicit than principles, they are not explicit enough with regard to actually implementing principles. Thus, Dumas & Redish introduce the notion of "local rules" (1999:58). Local rules provide clear, unambiguous and repeatable strategies for implementing the knowledge provided by principles. For example, if we use the principle "be consistent" and develop a guideline for it like "use the same formulations and constructions for sentences", we could produce the following rules:

- Always use active verb constructions when describing actions performed by the system
- Only refer to the product as X, never as Y or Z.
- The verb "run" must be used instead of "execute" or "call".
- Conditional sentences must take the form "If [condition], then [action]"
- Sentences must not exceed 20 words

A significant difference between guidelines and local rules is that while guidelines may conflict with each other in certain situations, rules **always** apply; they are absolute constants in the context where they apply.

Creating and compiling sets of guidelines and rules is complex and time consuming according to Dumas & Redish (1999:60). In this regard, they say that it is not always necessary to do so because there are numerous sources of guidelines and rules available. Such guides include the Microsoft Style Guide, the SAP Guide to Writing English, AECMA and so on. These publications provide a way of implementing various principles and guidelines without the effort or expense of drawing up complete sets of rules for each context.

Iconic Linkage

With the knowledge gained from the previous sections, the next step is to establish how to improve the interface, i.e. the user guide. Having discussed the concepts of principles, guidelines and rules it is now time to implement our understanding of human cognition, usability and cognetics. We can do this by selecting one guideline and implementing it in a practical context in order to gauge its overall effect on usability. The following sections introduce Iconic Linkage as one possible guideline and discuss its associated rules while seeking to elaborate on its practical implementation. The potential benefits of implementing this guideline will also be discussed.

A Definition of Iconic Linkage

Iconic Linkage (IL) refers to the use of isomorphic constructions to express what is, essentially, the same information. So, where the same information is expressed more than once in a text, the exact same textual formulation or construction is used. This is in contrast to the use of slightly different formulations which is employed in order to prevent what is commonly seen as repetition.

More specifically, in the case of translated texts, IL is the repetition or reuse of target language translations for source language sentences which have the same meaning but different surface properties. In other words, sentences which are semantically identical but which are non-isomorphic are translated using the same target language sentence or construction.

Origins of Iconic Linkage

The term Iconic Linkage was coined by House (1981:55) to refer to instances of structural similarity in adjacent sentences. She defines IL as occurring when two or more adjacent sentences in a text "cohere because they are, at the surface level, isomorphic". This phenomenon is quite similar to what technical writers call "parallelism" (see Chapter 2). Parallelism is a phenomenon which is widely recognised as a desirable feature of sentence structure (D'Agenais & Carruthers 1985:104; Mancuso 1990:231; White 1996:182). Essentially, parallelism means that parts of a sentence which are similar, or parallel, in meaning should be parallel in structure.

Parallel constructions can also be described as instances where two or more groups of words share the same pattern (White 1996:182). Thus, we can see that parallelism can occur on both a sentence level and on a sub-sentence level. The following sentences 1a and 1b illustrate parallelism.

1a: If you want to open a file, click Open.

1b: If you want to close a file, click Close.

Table 1: Example of parallelism between two sentences

When there is a lack of parallelism (for example in 2a and 2b) some of the grammatical elements in a sentence do not balance with the other elements in the sentence or another sentence. Consequently, the clarity and readability of a section of text are adversely affected. What makes this undesirable, apart from potential grammatical errors, is that it distracts the reader and prevents the message from being read quickly and clearly (Mancuso 1990:232).

2a: If you want to open a file, click Open.

2b: The Close button must be pressed when you want to close a file.

Table 2: Example of Two Sentences which do not have Parallel Structures

Parallelism is not just important in avoiding grammatical and comprehension problems, it is also very useful in reinforcing ideas and learning. The grammatical symmetry of parallelisms helps readers remember information more easily (White 1996:183).

Where my definition of IL differs from both House's definition and parallelism is that parallelism and House's definition deal with localised instances of structural similarity. Both deal with isolated pieces of text at

particular locations, e.g. a sentence or list. Instead, IL as used here is an active strategy which is used throughout a text. Indeed, instances of IL can be separated by large stretches of text. In addition, rather than being restricted to individual phrases or sentences, IL can manifest itself in entire paragraphs or longer stretches of text.

In contrast to House's definition, IL is actively introduced into a text, rather than being a naturally occurring feature of a text, i.e. a feature of the text when it was **first** produced.

Latent and Introduced Iconic Linkage

Before embarking on a more detailed discussion of Iconic Linkage, it is important to differentiate between the two principal types of Iconic Linkage: *Latent* and *Introduced*.

Latent Iconic Linkage refers to isomorphic, semantically identical sentences which occur "naturally" in a source text. These instances of Iconic Linkage form part of the text as it was originally written. Frequently, such instances of Iconic Linkage are destroyed during subsequent processes such as editing or translation. This can occur for a variety of reasons. With regard to the translation process, translators may not always remember how they dealt with a previous instance of a sentence and will provide a slightly different rendering. While this problem may be lessened though the use of translation memory (TM) tools such as Trados *Translator's Workbench* or Atril *Déjà vu*, not all translators have access to such tools. Another reason for the loss of latent Iconic Linkage both during translation and during editing is that repetition can be regarded as anathema to good writing. While this may be the case with creative and other forms of writing, repetition (or consistency) is actually desirable in technical documents in general and user guides specifically. As such, all instances of latent Iconic Linkage represent naturally occurring "quality components" within the text and should be maintained during translation and editing.

Introduced Iconic Linkage refers to instances of Iconic Linkage which are added to a text *during editing or translation*. If for example, two or more non-isomorphic but semantically identical sentences (or parts of sentences) are encountered by an editor or translator, using the same target text construction to render each source sentence will introduce Iconic Linkage into the text.

Types of Iconic Linkage

There are two fundamental types of Iconic Linkage: *Full and Partial*. Each type is determined by the extent of repetition within a particular sentence. The following paragraphs describe the nature of each type.

Full Iconic Linkage

Full Iconic Linkage refers to complete, discrete and relatively independent units of text such as sentences, headings, table cells etc. which are semantically identical and which are also isomorphic. Rather than writing the same information using different constructions or formulations, the same unit is repeated at different points in the text. Table 3 illustrates examples of full Iconic Linkage.

Matching Paragraphs

Full Iconic Linkage, i.e. identical sentences may be combined to form larger sections of isomorphic, semantically identical text. Thus, several instances of full Iconic Linkage can form iconically linked paragraphs which can then in turn be combined to form even longer stretches of iconically linked text.

Partial Iconic Linkage

Partial Iconic Linkage refers to *parts* of a unit that are identical – this is usually because there are certain factual differences which mean the units are not complete semantic matches. It can also be because one unit communicates more information than the other. Ideally this should not happen because best practice in technical writing holds that a sentence should only communicate one idea at a time. This also complies with the principles of good design set out by cognitive design principles to reduce STM load (see Chapter 3). Table 4 illustrates some examples of partial Iconic Linkage.

Examples of Latent Iconic Linkage

The following tables illustrate different forms of latent Iconic Linkage: *Full Iconic Linkage* and *Partial Iconic Linkage*. These examples are drawn from a selection of user guides which I have either translated or edited in my capacity as a freelance translator.

Examples of Full Iconic Linkage

Number	Sentence
3	If you make an error, press the Clear key.
6	The menus shown below are examples only and may not appear on your keypad.
2	The Save As dialog box will appear.
3	Use another universal controller.
4	The following table describes the XA-related flags.

Table 3: Examples of Latent Iconic Linkage – Full

Examples of Partial Iconic Linkage

Number	Sentence 1	Sentence 2
2	You can only *view* the date and time if your keypad has been programmed to do so.	You can only *change* the date and time if your keypad has been programmed to do so.
6	You can only *XXXX* if your keypad has been programmed to do so. (x6)	
2	If the *globals have* been saved as a UCC4 file before, select Save UCC4 from the File menu.	If the *file has* been saved as a UC16 file before, select Save UC16 from the File menu.
2	This will save *them in a file* automatically with the extension .cmn.	This will save *the file* automatically with the extension .stg.
2	*ON* This is the time the Schedule *starts*.	*OFF* This is the time the Schedule *ends*.
2	Up to *8* datalogs may be stored by a *UC12*.	Up to *16* datalogs may be stored by a *UC16PG or UC16IP*.

Table 4: Examples of Latent Iconic Linkage – Partial

Examples of Introduced Iconic Linkage

As mentioned previously, Iconic Linkage can be introduced into a text either during initial writing, editing or translation. The following examples, based on the user guide for a software package called *DigiLog* (see Chapter 5),

illustrate the introduction of Iconic Linkage in a monolingual context, i.e. production or editing. These examples present sentences which are all semantically identical. Each sentence in each group can be used to replace the other two sentences in each group.

Examples of Introduced Full Iconic Linkage

1a	Enable *QuicKeys* by clicking *Use QuicKeys*.
1b	*QuicKeys* are enabled by clicking *Use QuicKeys*.
1c	To enable QuicKeys, click Use Quickeys.
2a	If no checkmark appears before the *Use QuicKeys* option, the *QuicKey* function is disabled.
2b	If there is no checkmark in front of the *Use QuicKeys* option, the *QuicKey* function is disabled.
2c	*QuicKeys* are disabled when there is no checkmark in front of the *Use QuicKeys* option.
3a	Select the *Exit* option to close the *DigiLog* program.
3b	Click *Exit* to terminate *DigiLog*.
3c	To close *DigiLog*, click *Exit*.

Table 5: Introduced Iconic Linkage in a Monolingual Context

The sentences shown in Table 6 (from the author's personal translation archive) provide examples of Iconic Linkage introduced during translation and are taken from a user guide for a food packaging machine and its accompanying software. They illustrate how non-isomorphic but semantically identical sentences in a source text can be rendered with Iconic Linkage in the target text. The various sentences have been analysed using the Trados Translator's Workbench translation memory tool to illustrate the apparent match between sentences based on surface characteristics.

Original Instance	Wechsel in den Programmier-Modus (PROG-Modus).
Second Instance	Wechseln Sie in den PROG-Modus.
Translation	Switch the machine into program mode (PROG Mode).
(The two sentences are a 54% surface match but actually mean the same thing)	
Original Instance	Maschine ist im Arbeits-Modus.
Second Instance	Die Maschine befindet sich im Arbeits-Modus.
Translation	The machine is now in Work Mode.
(The two sentences are a 67% surface match but actually mean the same thing)	
Original Instance	Wollen Sie ohne Begasung arbeiten, stellen Sie den Parameter-Wert Gas/Restvakuum auf den selben Wert ein, wie den Parameter-Wert Vakuum.
Second Instance	Für das Arbeiten ohne Begasung müssen Sie den Parameter-Wert Gas/Restvakuum auf den gleichen Wert einstellen wie den Parameter-Wert Vakuum.
Translation	If you want to deactivate the gassing function, use the same value for Gas/Residual Vacuum and Vacuum.
(The two sentences are a 72% surface match but actually mean the same thing)	

Table 6: Examples of Introduced Iconic Linkage – Full

Examples of Introduced Partial Iconic Linkage

The sentences shown in Table 7 provide examples of introduced partial iconic linkage. These examples differ from each other in terms of certain material facts, e.g. in each case a different button must be pressed or a different system state occurs.

Original Instance	Betätigen Sie die *PROG*-Taste.
Second Instance	Drücken Sie nun die *ENTER*-Taste.
Translation	Press the *XXXX* button.
(The two sentences are a XX% surface match)	
Original Instance	Die Maschine bleibt im *PROG*-Modus.
Second Instance	Die Maschine befindet sich im *Arbeits*-Modus.
Translation	The machine is now in *XXXX* Mode.
(The two sentences are a 67% surface match)	

Table 7: Examples of Introduced Iconic Linkage – Partial

Iconic Linkage as a Cognitive Strategy

As discussed earlier, principles are general goals or objectives which may give rise to several guidelines and local rules. So for example, the principle which states that a user's STM capacity of 7±2 chunks of information should not be overburdened might give rise to guidelines to the effect that sentences should be kept short and contain simple and familiar words (which can be grouped into larger chunks of information, thereby reducing the number of chunks the reader has to retain).

However, these guidelines can also prove effective in implementing other, perhaps completely different principles. For instance, the idea of using simple and familiar words may also be used as a way of minimising users' memory load because people find it easier to recall the meaning of words which are used frequently. Similarly, while keeping sentences short may reduce STM load, it can also reduce the possibility of the reader becoming distracted by some other stimulus. This idea is related to the principles of saying only what is necessary, presenting information in a logical

way, and preventing errors by limiting the number of available options or choices. As Kellogg puts it

> …adopting cognitive strategies that circumvent attentional overload may be one key to optimum writing performance (Kellogg 1988:355-6)

If we relate Iconic Linkage to the concept of principles, guidelines and rules, we can say that Iconic Linkage is a guideline which states "always present information which is repeated throughout a text using exactly the same formulation of words and structures". This guideline itself draws on several principles such as consistency, reducing STM load etc., but gives rise to manifold rules which can consist of various efforts to regulate the ways in which information can be formulated and presented in a text.

The Benefits of Iconic Linkage

The idea of implementing Iconic Linkage in a user guide draws on several areas of cognitive psychology and represents a way of implementing a variety of principles. On the basis of the areas examined in Chapters 2 and 3 we can see that, in theory at least, Iconic Linkage presents the following benefits in the production of user guides.

Reduction of Short Term Memory Load

Iconic linkage reduces the demands placed on short-term memory by helping to chunk information into meaningful and manageable units. Not only does this help maximise on STM capacity but it also reduces the amount of processing needed and speeds up the retrieval of information.

Habit Formation

Iconic Linkage facilitates the formation of good habits. Raskin (2000:20) maintains that one obstacle to the formation of habits is the provision of multiple methods of doing something because it shifts the attention away from the task to the process of choosing the method. By ensuring that information is phrased in a standard, uniform way, we eliminate the need for readers to decode the same information all over again. Thus, users concentrate more on the task at hand and not on the interface (user guide).

Subconscious Learning & Perceptual Learning

In a way similar to habit formation, the repetition of information phrased in precisely the same way takes advantage of latent learning (see Chapter 3) whereby users learn information without realising it or intending to.

As we discussed earlier in Chapter 3, perceptual learning is the way we know **how** to perceive sensory information each time we encounter it. This is important for Iconic Linkage in that we can change the response of users to sensory perception. For example, where the sight of a particular piece of text may result in confusion, over time it will provide reassurance, information etc. for readers.

Accessibility of Information

By consistently using the same constructions to present information, the text assumes a certain visual or iconic consistency with each visual representation being associated with the information it conveys in image memory. This in turn allows Iconic Linkage to take advantage of image memory (see also Faulkner 1998:4) which reduces the processing load and speeds up retrieval of information by using recognition as a precursor to recall; recognition allows us to determine which information we have already seen and need to recall and which information we have yet to learn.

Retention of Information

Since repetition aids habit formation (see Chapter 3), Iconic Linkage, by virtue of the fact that it repeats textual constructions, increases retention of information. Furthermore, borrowing from the theory behind parallelisms, the grammatical symmetry of parallelisms also helps readers remember information more easily and it reduces confusion while improving clarity and readability (cf. Mancuso 1990:231; White 1996:182).

Problem-Solving

Since Iconic Linkage reduces the cognitive resources needed to decode and access information, it reduces the amount of problem-solving required to comprehend a piece of text. This is because the "problem", i.e. a sentence, has already been solved and so each time the sentence is encountered the solution is reinforced rather than created again. This has implications for the minimax principle introduced by Levý (1967) and incorporated into Gutt's relevance approach (2000:8).

Improvement of Consistency & Predictability

It is clear that using the same phrases and constructions to present the same information improves the consistency and predictability of the text. With no major surprises lying in wait for users, they can proceed through the text in relative comfort.

Reduction of Interference

By making texts more predictable and consistent, Iconic Linkage reduces interference between tasks. When we perform two tasks simultaneously, e.g. learning new information and using a piece of software, our performance in both will suffer, regardless of how automated one task (such as reading) has become. The more predictable, automatic and unconscious a task becomes, the less likely it will degrade or compete with other tasks for resources. And so, rather than having the process of reading the user guide competing with the task of using the software for cognitive resources, we can automate the reading process even more. As a result we can reduce the cognitive demands reading makes on the user and free up more resources for the process of using the software.

Potential Criticisms of Iconic Linkage

While the benefits mentioned above are clearly quite promising, it is conceivable that some would question the merits of such an approach. Perhaps one of the most significant criticisms that could be levelled at the notion of Iconic Linkage is that because it renders different sentences, which may have slightly different nuances, using identical translations, it will "water down" the message in the source text by removing various nuances of meaning. Under different circumstances, this could be a valid criticism. However, in technical communication, as we established in Chapter 2, our sole objective is to convey information which will allow readers to either perform some task or to learn something. Technical texts are not intended to entertain or impress, nor are they supposed to demonstrate any literary tendencies. Consequently, provided our translation conveys the essential information and permits the reader to perform the task at hand, any omission of cosmetic or stylistic features is perfectly acceptable.

Similarly, it could be argued that Iconic Linkage exacerbates what Levý (1969:110-111) describes as the "lexical impoverishment" brought about by translation. This, he says, comes about because translators tend to use more general vocabulary than original language writers. By implementing Iconic

Linkage, we could be accused of making this situation more acute. But we should ask ourselves whether this is really a bad thing. After all, the aim of technical communication is to communicate information with a maximum of clarity and a minimum of ambiguity. In any case, technical documents tend to restrict the language used in them to minimise misunderstanding. So if the use of Iconic Linkage does lead to lexical impoverishment, perhaps this is another benefit rather than a criticism.

A final criticism which will be raised here is that Iconic Linkage introduces a level of monotony into a text which may make the text boring at best. It could be argued that the repetitive nature of a text as a result of Iconic Linkage may even distract or alienate readers. In response to this, I would refer again to the fact that technical documents are not intended to be amusing or entertaining. Their sole purpose is to communicate technical information clearly and efficiently. However, he idea that monotony may in some way impair readers' use of a text is worthy of further investigation. This will be examined in the following chapter as part of an empirical study of usability.

Conclusions

Building upon the insight into cognition provided in Chapter 3, this chapter examined usability from the point of view of interfaces. In the context of human-computer interaction, an interface can be defined as anything that acts as an intermediary between a user and the inner workings and processes of a computer or piece of software. Having described usability in Chapter 2 as ensuring that an interface takes into account the cognitive abilities and limitations of humans, this chapter turned the notion of usability into a series of definite design goals and objectives which ensure users can work effectively, efficiently and with ease. These goals codify those aspects of human cognition where active strategies can be implemented to improve the interaction.

Cognetics, or cognitive engineering, was introduced as a discipline where interfaces are engineered or designed to capitalise on human abilities and to compensate for human limitations. The first step in any attempt to improve the usability of a user guide is to set goals. With a clear understanding of what it is we need to achieve when producing user guides, the next step was to establish how to actually improve the interface. The chapter then discussed the concepts of principles, guidelines and rules and provided examples of each of these concepts.

Iconic Linkage (IL) was then introduced as one possible guideline which allows translators to improve the usability of a user guide. We saw that IL can occur naturally within a text (i.e. when the text is first produced) or it can be actively introduced into a text during the translation process.

We also looked at the extent to which partial or full IL can occur between two or more sentences. Where only parts of the information payload in sentences are matched and this information is phrased identically, the IL is said to be *partial*. Where entire sentences are semantically identical, both can be written using the same structures and formulations. Such sentences are semantic matches and represent examples of *full Iconic Linkage*. Where several instances of full IL occur consecutively, they can represent *matching paragraphs* where the whole paragraph presents the same information in exactly the same way as in another part of the text. Again, several matching paragraphs occurring in succession result in larger sections of Iconic Linkage. As a translation strategy, IL represents a bridge between the various principles outlined in preceding chapters and the numerous technical writing rules.

The following chapter has two primary aims: firstly, it will illustrate how Iconic Linkage can be implemented in user guides during the translation process. Secondly, it will discuss a variety of methods and approaches which can be used to evaluate the usability of user guides in general before developing a model for assessing the effect of Iconic Linkage on document usability.

Chapter 5
Assessing Usability

The preceding chapters examined software user guides with a view to improving their usability. Iconic Linkage was proposed as one method of improving the usability of user guides. However, a crucial part of usability engineering is being able to assess the effects of any strategy aimed at improving usability. This chapter deals primarily with an empirical study conducted to test the effect of Iconic Linkage in a software user guide under conditions representative of real-world usage. A crucial part of such an undertaking is examining the various methods and procedures commonly used in studies of this nature. The chapter begins by examining a range of data collection methods used in usability studies and proceeds to select those which are compatible with assessing the usability of a printed user guide. To gain a better insight into usability testing involving users, we will also look at a number of previous usability studies.

We will look at how a model for conducting the empirical experiments was developed including hypothesis, evaluation criteria, methods, materials and data analysis models. To conclude, this chapter will discuss and evaluate the findings of the empirical study.

Empirical Usability Evaluation

Regardless of the time and effort spent on engineering usability into an interface using the variety of methods outlined in previous chapters, the only true way of establishing whether the interface is indeed usable is to conduct some form of usability evaluation. There are essentially two types of evaluation – formative evaluation and summative evaluation. The type of evaluation used depends on when the evaluation is to be carried out as well as what it is hoped the evaluation will achieve.

Formative evaluation takes place during the development process in order to detect potential problems before the design is actually implemented.

The aim here is to improve the usability of the final interface (Preece 1993:108; Preece 1994:603).

Summative evaluation, in contrast, is carried out on the finished interface. The aim here is to determine the level of usability of the interface so that judgements can be made as to the overall usability and quality of the interface (Preece 1994:103). Summative evaluation is used to ensure that "the final product is up to standard, not to shape the design and development processes" (Landauer 1997:204).

For the purposes of assessing a finished translation, we are interested in assessing the overall level of usability of the final product, i.e. the user guide. For this reason, we will restrict ourselves to an examination of usability evaluation from a summative point of view.

Historical Problems in Document Usability Research

A review of the literature on usability evaluation reveals a range of publications concerned with improving the usability of software documentation. However, much of the research undertaken in recent years is less than satisfactory for our purposes here for a number of reasons.

First of all, the term *documentation* as used by several authors proves problematic, not only in terms of the aims of this book, but also in terms of what happens in an industrial context. Documentation is frequently defined in an extremely broad sense as including practically anything that involves some form of text. Indeed, Mamone (2000:26) defines documentation as

> …user manuals, online help, design features and specifications, source code comments, test plans and test reports, and anything else written that explains how something should work or be used.

While this is perfectly acceptable and indeed suitably comprehensive in many ways, it is problematic from the point of view of usability evaluation. Mamone's assertion (*ibid.*) that documentation can come in "hard or soft form" fails to take into account the fact that while technology and economics have pushed the online delivery of information, online and hardcopy documentation have specific strengths and weaknesses (Smart & Whiting 1994:7). As such, they cannot be assessed using identical sets of criteria.

Indeed, the all-inclusive definition of documentation includes online help. In addition to the obvious textual issues involved in using help and

the fact that information is presented as short, **independent** units, the diversity and sophistication of help systems mean that there is quite a different set of considerations (such as navigation models, resolution, display speeds, download times, delivery options etc.) to be borne in mind when evaluating their usability in comparison with that of hardcopy documentation. So, while Mamone provides a useful overview of usability test procedures, the failure to identify the differences and the similarities between hardcopy documentation and online help compromises his approach. But it is not just Mamone who fails to make this distinction, Prescott & Crichton (1999), Harrison & Mancey (1998), Simpson (1990) and Mehlenbacher (1993), to name a few, either concentrate on online texts or group online texts together with print documentation.

Although the practical issues relating to the evaluation of online documentation mean that online and print documentation cannot reasonably be assessed using a single theoretical framework, other factors justify the separate treatment of the two types of text. By grouping print and online texts together, print documentation risks being regarded as a "low-tech" form of documentation which can be assessed with just a subset of the online evaluation paradigm. Admittedly, online documentation is increasingly regarded as an integral part of products which speeds up the dissemination of information and which – if designed well – can allow users to access information more quickly. Nevertheless, a large proportion of users prefer the "book" format. According to Smart & Whiting (1994:7) "some information is best accessed from a printed form." Such information includes trouble shooting, lengthy conceptual descriptions or material where annotation is essential. With this in mind, we can see the need to examine documentation purely from a print point of view.

Approaches to Empirical Evaluation

There are several approaches and methods for assessing usability which can be used. We can group these methods into two broad categories: *analytical* and *empirical* (Faulkner 1998:113).

Analytical evaluation, according to Preece (1993:109), uses formal or semi-formal methods of describing the interface in order to predict user performance. These methods include such strategies as GOMS (*goals*, *operators*, *methods* and *selection rules*) and KAT (*knowledge analysis of tasks*). Analytical evaluation is primarily a formative approach and as such is of little significance here.

Empirical evaluation as described by Faulkner (*ibid.*) includes expert evaluation or heuristic analysis, observational evaluation, survey evaluation and cognitive walk-throughs (Hill 1995:120). Within this broad category of empirical evaluation, we can distinguish between *absolute* and *comparative* studies. Absolute experiments (for example, expert or heuristic evaluation) involve assessing the interface or system on the basis of predefined specifications, criteria and scores. Such standards might include, for example, Nielsen's heuristics (Nielsen & Molich 1990). Comparative experiments, on the other hand, involve assessing an interface and comparing it with some other alternative interface or version of the interface (Downton 1991:331; Faulkner 1998:113). For the purposes of this research, the comparative approach is preferable because we are concerned with determining whether Iconic Linkage can improve the usability of user guides in comparison to user guides where Iconic Linkage is not present.

We can further divide empirical evaluation into methods which involve users and methods which do not involve users. This distinction is important because Landauer (1995:281) maintains that, in contrast to analytical evaluation, expert evaluation and cognitive walkthroughs, testing involving users is the best way of evaluating usability. He says "only by studying real workers doing real jobs in real environments can we be sure that what we find out is truly relevant" (*ibid.*). With this in mind, the following pages will discuss only those empirical evaluation techniques which involve users, namely observational evaluation, survey evaluation and experimental evaluation.

A critical factor in any form of evaluation is the type of information the evaluation is supposed to provide: quantitative or qualitative. Quantitative data is numeric and is analysed using a range of statistical and mathematical methods. This, in some regards, makes it easier to process large amounts of data in order to provide statistical evidence. However, unlike qualitative data, it does not provide the detailed subjective data or opinions that give an insight into how users actually perform tasks using the interface. Qualitative evaluation does yield this information although the resulting data is more difficult to analyse in that it consists of a wide variety of diverse information which is often expressed in very different ways by users.

However, in the case of this study, we are interested in examining the way users work with the interface in a significant amount of detail. As we discussed in Chapter 4, usability is determined not only by how quickly or efficiently users work with an interface but also by users' opinions, satisfaction and attitudes to the interface. In addition to quantitative data such as how quickly users perform tasks, we are interested in finding out how easy users perceive the use of the interface to be. In addition, the decision as to

whether to gather quantitative or qualitative data has significant implications for the design of the experiment, the methods used and the number of subjects involved in the experiment.

In the following sections, we will examine the various methods for conducting usability evaluations with users in order to gather a combination of both qualitative and quantitative data.

Observational Evaluation

Observing users performing tasks they would normally perform as part of their work is one of the most useful ways of collecting data about what users do when they use an interface and how they react to the interface. Users can be observed in a specially built usability lab (Preece 1993:112; Schneiderman 1998:128-130) or in the normal environment where the users work. The latter scenario is often referred to as a field study (Preece 1994:602). Observation can take a variety of forms and there are several methods for recording the information. There are two fundamental forms of observation: *direct observation* and *indirect observation*.

Direct Observation

Direct observation involves users performing tasks while an observer is present in the same room. The observer watches the user and makes notes, times actions or performs some other function such as asking questions etc. While this approach is valued for its informality and immediacy, there are a number of problems associated with direct observation. The first and perhaps most important issue with regard to the validity of the test results is that the presence of an observer can have a significant impact on the users' performance. This may be attributed to users believing that their performance is under constant scrutiny and that they need to "impress" the observer. The overall result is that users may actually perform better under test conditions than they would under normal working conditions, simply because they are trying much harder, a phenomenon known as the *Hawthorne effect* (Preece 1994:617; Faulkner 1998:122).

A second problem with direct observation is that the quality of information gathered and indeed the completeness of the information relies on the ability of the observer to correctly interpret what is happening and then to write accurate and useful notes. Furthermore, if the observer misses something, there is no way of capturing the lost data or referring back to what happened – the experiments are single-pass occurrences and the only record

of events is the notes taken by the observer. Of course, using more than one observer may yield more comprehensive and complete data and also even counteract any possible biases associated with a single observer. However, if the presence of just one observer can affect a user's performance and distort the results of the test – the effects of several watchful observers could have disastrous consequences for the validity of the experiment.

Both of these problems, i.e. the Hawthorne effect and the unreliability/incompleteness of notes, can be counteracted through the use of indirect observation techniques (Faulkner 1998:122).

Indirect Observation

In contrast to direct observation where users perform tasks with an observer present in the same room, indirect observation involves users performing tasks without the presence of an observer. This type of observation generally incorporates some form of recording mechanism, be it in the form of audio, video or software recording or some combination of the three.

Audio Recording

Audio recording can be useful when combined with verbal protocols. It involves recording what a user says during the course of an experiment from general comments or spontaneous outbursts to verbalised thought processes in the case of think-aloud protocols (TAP). Audio recording and TAP are generally of greatest use in formative evaluations as the wealth of qualitative information can provide significant insight into the way the users interact with and perceive the interface. Particularly in the case of evaluations involving interactions with software, it can be difficult to match the audio recording with field notes on the events as there are no clues other than what users say and they may not always verbalise problem areas. It can be argued that TAP can place additional cognitive loads on subjects and can interfere with the way they perform tasks (see the discussion of attention in Chapter 3).

Video Recording

Video recording or logging counteracts both the problems of direct observation and also the problems of audio recording in that it does not require the observer to be present in the same room as the user and it provides a permanent record of the experiment while allowing the observer to see what the user did at a given point in the experiment. It can also be used in conjunction with some form of verbal protocol. By positioning cameras in

a variety of locations in the room where the test is being conducted it is possible to capture a wide range of data such as what the user types on the keyboard, what appears on the screen, whether the user refers to the user guide, as well as the user's body language or facial expressions. With modern video camera technology, it is also possible to record reasonably high quality audio as well, thereby negating the need to synchronise different types of data.

There are, however, certain problems associated with video logging. While this method provides valuable and comprehensive information, it only does this for as long as the user stays within the camera's field of view (Dix 1998:428). An obvious solution would be to position the camera closer to the user but then we risk undoing the benefits of indirect observation with the obtrusiveness of the camera. Conversely, hiding the camera and filming users surreptitiously raises certain ethical and legal issues and is to be avoided (Faulkner 1998:123; Dumas & Redish 1993:206). A simpler solution is presented by Dumas & Redish (1993:225). Rather than hoping that users stay in the correct position during the test, Dumas & Redish propose placing pieces of adhesive tape in an L-shape on the desk to indicate where the documentation must be placed. In addition to ensuring that the document stays within shot (if necessary) this strategy also constrains the users' movements and ensures that they stay in more or less the same position.

If recording events on the screen, careful positioning of cameras, the user, the screen and lighting is necessary unless there is some mechanism for connecting a second monitor (Dumas & Redish 1993:224, 384). In such a scenario, the second monitor could be placed in another room where the observer records the images on video. But the issue of data synchronisation can be quite problematic where more than one camera is used (Dix 1998:428; Preece 1994:619). At the very least some form of on-screen time code along with a mechanism for starting all of the cameras simultaneously will be necessary. This problem can be alleviated at least partially by through the use of software logging.

Software Logging

Software logging is where the computer system records the user's actions during the test. There are two basic forms of software logging: one which records *time-stamped keypresses* and one which records the user's *interactions* with the system.

Time-stamped keypress logging records each key a user presses along with the time it was pressed. Certain varieties of keypress loggers also record

system responses (Preece 1994:627) which means that error messages and dialog boxes can also be recorded.

Interaction logging tools operate in a similar manner except that they record the entire interaction in real-time. What makes interaction logging truly useful is that it allows the interaction to be replayed in real-time. This can illustrate additional information such as hesitations, visibly confused mouse movements and aborted attempts at tasks which might not be detected by keypress tools.

Apart from eliminating the need for a video camera to be recording the screen, the advantages of software logging are the fact that it is unobtrusive (although the same ethical questions apply as for video recording), it is at least partially automated and it provides a permanent record of the test. When used in conjunction with a video camera, it provides a comprehensive picture of the experiment from the point of view of the user, the interface and the interaction (Downton 1991:333). The only major drawback with this method is that it may require large amounts of storage space to deal with the frequently huge volumes of data and there may be synchronisation issues between the screen recording and the video recording (Preece 1994:266-267; Preece 1993:113; Dix 1998:428). However, this is offset by the fact that there is no need to analyse tapes from a second camera aimed at the screen.

Interactive Observation

Interactive observation is a type of indirect observation where the part of the system or computer is played by a member of the evaluation team. Commonly known as the "Wizard of Oz" (Faulkner 1998:122), this approach makes users think that they are using a real system, but in reality all of the system's responses and actions are performed by a human. This method is effective in that it does not require a fully functional version of the system and it can be implemented reasonably cheaply in comparison to the expense of producing a fully-functioning system. However, this approach is formative and is more suited to situations where an actual software system is being produced. The effort required to create an interface, in terms of design and labour coupled with the additional staff requirements to conduct the experiment make this approach difficult to implement. In any case, this approach is of limited applicability to documentation as it is aimed at investigating the way the software interface works.

Verbal Protocols

Verbal protocols are spoken records of users' comments, observations, exclamations and other information which may arise during the course of an experiment. One particular variety of verbal protocol is the think-aloud protocol which involves users saying what they are thinking, feeling, planning etc. as they perform tasks and use the interface. This can provide a valuable insight into what users want to do, what they think they are doing and what their responses are when something unexpected occurs. It is also possible to gain an insight into how users remember commands, plan and execute tasks and how they recover from errors (Preece 1993:113). Verbal protocols of this type are generally used in conjunction with audio or video recording (Preece 1994:621).

Although think-aloud protocols (TAP) are ideal for formative usability testing where the wealth of qualitative data they can provide is extremely useful in understanding the nature of the interaction, in the case of purely *summative* evaluations where, for example, the speed at which users work is being measured, TAP is less applicable, chiefly because summative evaluations require quantitative data. It can also be argued that TAP may hinder evaluation rather than aid it. This can be attributed to a number of factors. Firstly, as we discussed in Chapter 3 the human cognitive system can realistically deal only with one response to a stimulus at a time even though it can process several inputs or stimuli. Indeed, the process of dividing one's attention equally between two tasks is unreliable at best, but extremely difficult when performing two complex tasks such as problem-solving and verbalising thoughts where high levels of accuracy are required for both. Similarly, it is also held that the very act of putting into words what it is a user is doing will affect the way the user performs the task (Dix 1998:427; Downton 1991:334). While many agree that this double-tasking will degrade performance on both tasks, there is some conflicting evidence that the think-aloud protocol may actually improve performance of the task. If this were proven to be true, it could be because the verbalisation process focuses a user's mind on the task and helps users rationalise the task better. Nevertheless, the additional strain of performing two complex and demanding tasks such as putting thoughts into words can result in lower performance and some users will simply be unable to verbalise their thoughts (Preece 1994:622).

There is also the problem of silence caused by the fact that users are either unaccustomed to thinking out loud or because all of their concentration is being devoted to performing the task; some users may simply forget to speak. This problem is discussed by Dumas & Redish (1993:278–281)

who point out that while some people have no problems whatsoever in producing an "unedited stream of consciousness", others either mumble or do not speak. The authors make the interesting point that users need to be taught how to think out loud and that they may need to be reminded to do so. This in itself can become a source of stress for users who may already feel pressurised as a result of the tasks (Preece 1994:622).

In view of these problems, *retrospective* or *post-event protocols* are sometimes used to elicit verbal data from users. Instead of commenting

on their actions while they perform them, users are shown a video of the experiment and are asked to comment on theiractivities. This approach produces different results in terms of the type of information users provide. According to Preece, users tend to rationalise or interpret their actions or even justify them (Preece 1994:623). Rather than simply stating what they were doing or thinking, users tend to explain what they are doing and why.

Dumas & Redish, however, do make the point that retrospective protocols frequently yield more suggestions as to how to improve the interface as compared to think-aloud protocols (Dumas & Redish 1993:279). However, we are not interested in using the evaluation to improve quality, merely to assess it.

It is clear from the preceding paragraphs that observational methods are extremely useful in gathering comprehensive information on the way users work with an interface. While these methods produce large volumes of data which can be more difficult to analyse, the sheer detail and insight they provide more than compensates for this (Preece 1993:119). By using indirect observation we avoid such negative effects as the Hawthorne effect and we are provided with a permanent record of the experiment. In order to ensure that participants' task performance during the experiments is as representative of real-life as possible, think-aloud protocols are to be avoided as they can affect the way tasks are performed. Retrospective protocols are of limited use and are really only of benefit when the purpose of the evaluation is improvement rather than quantification; in the case of this study, we are concerned with the latter.

Survey Methods

In the previous sections detailing the various observational methods, we examined evaluation methods which provide us with objective data about how users work with an interface. From this data we can see exactly how well users perform tasks using the system and also where any problems are.

This information is, without doubt, of enormous value but it is not enough on its own. To really understand whether an interface meets users' requirements, we need to elicit subjective information from users to illustrate their attitudes to and perceptions of the system (Dix 1998:431). Indeed, as Preece (1994:628) points out, users' opinions can affect the design of an interface while their attitudes affect the acceptance of the interface in the workplace. In short, if users do not like an interface, they will not use it unless they absolutely have to.

User survey methods – also known as query techniques or subjective assessment – make it possible to target large numbers of users to obtain their opinions directly and to highlight problems which may not have been anticipated by designers or evaluators (Dix *ibid.*). There are two main types of survey technique: *interviews* and *questionnaires*.

Interviews

The process of interviewing users regarding their experiences with an interface is a direct and structured way of gathering subjective information (Dix 1998:432). Interviews can generally take one of three forms: *structured interviews*, *flexible interviews* and *semi-structured interviews*.

Structured Interviews

In a *structured interview*, the interviewer uses a fixed and predetermined series of questions which are asked in strict order. This approach allows for close control of the type of data gathered and makes it easier to analyse. By limiting the possibilities for tangential discussions and comments, structured interviews ensure that interviewers are not distracted from the true aim of the evaluation and that the desired information is obtained. This type of interview is generally easier to conduct and it is easier to analyse the results (Hill 1995:130).

However, due to their lack of flexibility, structured interviews do not allow interviewers to follow up new lines of enquiry or discover new information. Nevertheless, the nature of this study means that the information gathered from structured interviews, while of potential interest, is not essential for the purposes of the evaluation. Here we are more concerned with user attitudes and opinions than with their observations and suggestions as to how to improve the interface.

Flexible Interviews

Flexible interviews, on the other hand, have no set questions, only a number of set topics to guide the interviewer. With no set questions, the interviewer is free to follow any line of questioning that is of interest. This method provides much more information than the structured interview but this also means that the data will be more difficult to analyse. In addition, this type of interview requires experienced interviewers to put interviewees at ease and a considerable amount of effort to analyse the data (Preece 1994: 628-629). In addition, the sheer volume of information generated may prove problematic and it is also possible that interesting new lines of questioning may result in key information being omitted or forgotten.

Semi-Structured Interviews

The *semi-structured interview* is a hybrid technique which lies between structured and flexible interviews. This type of interview features a series of questions which can be asked in any order, combined or even omitted as necessary. This gives the interviewer a significant level of freedom to explore new lines of enquiry which may emerge during the interview while still ensuring that the required information is gathered. However, this type of interview, like the flexible interview requires experienced interviewers and a considerable amount of effort to analyse the data (Preece 1994: 628-629). The availability of experienced interviewers is, therefore, a major factor when considering this approach. As Downton (1991:337) points out, interviewing for surveys is a skill in itself which requires training. This can make flexible interviews impractical in many cases. In our case, the problems of finding skilled interviewers coupled with the intensive resource requirements (e.g. time, finance, processing etc.) make interviews unsuitable for our purposes.

Questionnaires

Questionnaires are less flexible than interviews but they take less time to administer, can reach a larger audience and the resulting data is easier to analyse (Dix 1998:432; Hill 1995:130; Downton 1991:334). However, a questionnaire, because of its static nature, needs to be carefully worded to ensure that participants understand the questions fully and provide the desired information.

Questionnaires can be either self-administered or interviewer-administered. Self-administered questionnaires are completed by users alone and are frequently posted out to users. While the staff resources are generally less than for interviewer-administered questionnaires, this particular

benefit is outweighed by the fact that clarification cannot be given to users. Consequently, there may be instances where users either do not answer questions or misunderstand them and give a false or misleading answer.

An additional problem with self-administered questionnaires is the frequently poor response rates (Faulkner 1998:118). Downton (1991:335) cites a response rate of less than 40% for postal questionnaires. This can result in extreme bias of results because of small, unrepresentative subject populations (*ibid.*).

Self-administered questionnaires also place greater pressure on the designers of the questionnaire to produce absolutely clear and unambiguous questions in order to ensure that users understand all of the questions. In order to do this a continual process of design, evaluation and redesign needs to be carried out until there is absolutely no room for confusion (Faulkner 1998:117). This would undoubtedly result in a lengthy and time-consuming process which would place additional demands on time, finances and other resources.

Interviewer-administered questionnaires involve an interviewer asking the questions and completing the questionnaire with the user's responses. Although this method requires the availability of an interviewer, the interviewer does not need the same level of skill or experience as for interviews. Furthermore, interviewer-administered questionnaires make it possible to better control the data gathering process (Downton 1991:335) and any confusion as regards unclear questions can be clarified immediately. In addition, the use of interviewers ensures that the poor response rates associated with self-administered questionnaires are avoided.

Types of Questions

There are three basic types of questions that can be used in a questionnaire: *factual*, *opinion* and *attitude*.

Factual questions, as the name suggests, ask about facts and information which is observable and public but which would be too time consuming or inconvenient to obtain any other way (Kirakowski 2000). Examples of such questions might include asking users which software packages they have experience of, how frequently people use a particular piece of software, how long people have been using a PC on average etc.

Opinion questions ask respondents to say what they think about something. Such questions aim to determine how popular something is or whether respondents like something or prefer one thing over another (*ibid.*).

Attitude questions aim to uncover a respondent's "internal response to events and situations in their lives" (*ibid.*). Such questions seek to find out what users' attitudes are to working with an interface. Using questions like this we can categorise users' attitudes to working with a product as follows:

- users' feelings of being efficient
- the degree to which the users like the system or interface
- how helpful the users feel the system or interface is

- the extent to which users feel in control of the interaction
- the extent to which users feel that they can learn more about the system by using it.

Presentation of Questions

In addition to the broad types of questions outlined above, there are two fundamental styles of question which can be used to elicit the desired information: *open* and *closed* questions. Open questions ask users to provide answers in their own words. Closed questions ask users to select their answer from a predefined list of options.

Both styles of question have their own distinct advantages and disadvantages. For example, open questions provide a wealth of information covering a broad range of issues but they are difficult to analyse on account of the sheer volume of data produced and the variations in the style and content of responses (Dix 1998:433). Faulkner (1998:117) shares this opinion and says "the problem with open questions is that they can produce too much data which is not easily analysed".

Closed questions, on the other hand, are generally easier to analyse than open questions and they allow evaluators to focus on specific data which can be compared against other data. Furthermore, the data obtained from closed questions is more predictable and requires less interpretation (Downton 1991:336). However, such questions need to be carefully phrased in order to elicit the precise information sought.

Types of Closed Questions

There is a range of ways in which closed questions can be structured in order to elicit a particular form of data. These range from simple checklists to more complex multi-point scales and semantic differential scales.

Checklists are the simplest form of closed question and they ask for basic responses to specific questions. This type of closed question is ideal for factual information such as which software packages respondents have used etc.

Multi-point or *scalar rated questions* ask respondents to rate a specific statement between two polar opposites. This approach is suitable for determining user opinions.

Likert scales are similar to multi-point scales but in this case, respondents are asked to indicate the extent to which they agree or disagree with a statement. According to Kirakowski (2000:8), it is necessary to prove "that each item of the questionnaire has a similar psychological 'weight' in the respondent's mind". In order to prove the reliability of such scales, some form of psychometric evaluation is necessary (*ibid.*).

Ranked order questions dispense with scales and ask respondents to number, in order of preference, a series of options. This approach is best used with a limited number of options, otherwise respondents may give arbitrary answers (Preece 1994:633).

Semantic differential questions are similar to Likert scales but they ask respondents to rate an interface on a scale between two diametrically opposed adjectives, e.g. clear-confusing or interesting-boring (Preece 1994:632).

Multiple-choice questions offer a range of explicit responses and respondents are asked to select either one or more of these options.

A concern with questions that involve some form of scale is the granularity or number of rating points provided to the respondent. This relates to the level of detail an evaluator wants to achieve in the collated data. If a broad general idea of preferences or opinions is required, a simple three-point scale is adequate. However, if a more detailed breakdown of preferences and opinions is required, a greater number of rating points can be used. However, simply adding more and more points on the scale can prove counter-productive. If we use a ten-point scale, it is likely that some respondents may find it difficult to differentiate between any two adjacent points on the scale with the result that they may arbitrarily pick points (Dix 1998:433).

On a related note, Kirakowski (2000:10) raises the question of whether to use an odd or even number of rating points on the scale. The reasoning behind this is that with odd-numbered scales where the central point corresponds to neutral opinions or undecided users, respondents may "go on auto-pilot" and select neutral points without giving any real thought to their choice. As a way of preventing this, an even numbers of option can be

used to "force" respondents "to go one way or another" (*ibid.*). This is, according to Kirakowski, unhelpful in that it does not cater for respondents who genuinely have no preference or strong opinions. It is also common for respondents to randomly pick alternate options from the two middle options.

Developing a Questionnaire

Kirakowski (2000:6) maintains that developing a questionnaire requires "a lot of time, patience and resources". In addition, considerable knowledge of psychological measurement and statistics is essential. Kirakowski also maintains that unless all of these prerequisites are met, serious questions arise as to the validity and reliability of the questionnaire. As a result, he recommends that a pre-designed, pre-tested questionnaire should be used instead. In view of this, we will examine a number of commonly available usability questionnaires and select a suitable one. These range from freely available, public-domain versions to sophisticated commercial varieties.

A number of the available models are simply unsuitable for testing documentation, e.g. *WAMMI*[1] which is designed for testing websites. Others, such as *IsoNorm*[2] (developed by Jochim Pümper on the basis of ISO9241-10) are only available in languages other than English which could not be understood by participants.

Of the remaining models, the following criteria can be used to select the most appropriate one:

- cost
- reliability
- data validation & standardisation
- flexibility/customisation

The first such model is the *Computer System Usability Questionnaire* or CSUQ[3]. This questionnaire is available free of charge and incorporates psychometric reliability properties. However, the test is run using a web page which is submitted to another server for processing. While this removes the

---

[1] http://www.wammi.com

[2] http://www.sozialnetz-hessen.de/ergo-online/Software/Isonorm-Workshop.htm

[3] http://www.acm.org/~perlman/question.cgi?form=CSUQ

effort of processing data, it also leaves the test administrator without a permanent record of the answers provided. Another problem with this is the fact that questions are phrased in a way that could lead or prompt specific responses. Obviously, this can distort responses and lead to unreliable data. Furthermore, the questionnaire does not refer to documentation and it cannot be customised or modified.

Another questionnaire is the *Software Usability Measurement Inventory* or SUMI[4] developed by the *Human Factors Research Group* at University College Cork. This questionnaire overcomes the problems of CSUQ in that questions are worded in a variety of ways so as to avoid leading respondents. SUMI is supplemented with a data analysis and reporting service which eliminates the need for test administrators to perform complex calculations. However, samples of SUMI available on the Internet fail to include documentation. Above all, the high cost of SUMI can prove prohibitive.

The questionnaire selected for this study was the *Questionnaire for User Interaction Satisfaction* or QUIS[5] developed by Kent Norman and his research team at the *University of Maryland at College Park*. Using a 9-point Likert scale, QUIS features a section to determine overall levels of satisfaction and hierarchical sections which examine satisfaction with specific aspects of the interface (see Table 1).

• screen factors	• learning factors
• technical manuals	• multimedia
• virtual environments	• software installation
• terminology and system feedback	• system capabilities
• online tutorials	• voice recognition
• Internet access	

Table 1: QUIS Question Categories

What makes this model so attractive is that is can be modified and customised to suit the needs of a particular study. Thus, entire blocks of questions can be omitted. Just as importantly, all questions are phrased in a neutral

[4] http://sumi.ucc.ie

[5] http://www.cs.umd.edu/hcil/quis/

way which neither encourages nor discourages a particular response. On a practical level, QUIS can be administered as a web-based questionnaire (although this requires a specially configured web server) or as a paper-based questionnaire.

QUIS also comes with instructions for analysing and processing data (see Norman 2003). Indeed, this is reinforced by references to other research projects which have used QUIS. From a cost point of view, QUIS is comparable to other questionnaires such as SUMI but there is an option of obtaining a student site licence which is considerably more affordable.

Participants in a Usability Evaluation

As we have already discussed above, the aim of usability evaluation is to see how easy users find a system to use. It is obvious, therefore, that the participants in a usability evaluation should reflect the real users as accurately as possible. Consequently, finding and recruiting participants can be a complex process which must be planned carefully (Downton 1991:340). There are two key factors to be remembered when selecting participants for a usability evaluation: *who they are* and *how many of them are needed*.

Background of Participants

In asking who the participants are, we are concerned with those characteristics of the participants which may have some bearing on the evaluation. We need to ask questions such as what do they know? What skills and experience do they have? What are their ages, gender, background and level of education? But if there is to be any point in gathering this information, we must first know something about the real users of the system. In order to understand the characteristics of real users, we need user profiles so that we can select suitable participants to represent them in the usability evaluation (Dumas & Redish 1993:120).

Dumas & Redish (1993:122) propose a method of creating user profiles which involves usability experts working in conjunction with subject specialists or the actual users to define precise information relating to the users' backgrounds. Such information includes:

- work experience, e.g. job description, length of service, experience with particular tasks

- general computer experience, e.g. length of experience with specific types of applications, frequency of use etc.
- specific experience with operating systems, hardware (e.g. mouse or keyboard) etc.
- experience with this and similar products

Although Dumas & Redish do not explicitly say so, it is, of course, conceivable that this information could be gathered or supplemented by means of interviews or questionnaires. Alternatively, this information could be elicited, for example, from a manager or someone who supervises and recruits real users or possibly even someone who knows the technical skills of users, e.g. technical support engineers.

Dumas & Redish present a sample user profile which consists of five basic headings. Although this profile is intended for screening participants, it is, of course, also suitable for creating profiles of the real users.

- product name
- general characteristics of user population
- characteristics of users that are relevant to the test
- which user characteristics should all users have and how will you define them?
- which user characteristics will vary in the test and how will you define them?

We can expand this basic profile to produce a user profile questionnaire as shown in Appendix 1 Once this profile has been created for users, it is necessary to create a similar profile for each of the potential participants. With this information, selecting those participants who most closely match the real users is a relatively straight-forward process. The data obtained for participants can then be used to distribute participants with varying skills across a number of subgroups (Faulkner 1998:115). In our case, there will be two sub-groups: a control group and an experimental group. This is important because if one group has more participants with a particular type of skill relevant to the test than another group, the results may be distorted (*ibid.*).

Numbers of Participants

A critical element determining the success of any experiment in terms of validity and accuracy is the numbers of participants involved. Quite simply, there have to be enough participants in the test to ensure that the results are truly representative of the real users as a whole and that the results are not just the idiosyncratic behaviour of participants (Dumas & Redish 1993:128). Nevertheless, it would be impractical and indeed extremely difficult to use vast numbers of participants. Obviously some form of compromise is needed.

Nielsen (2001) maintains that for the purposes of usability testing in industry, 3-5 participants are generally sufficient in order to gain an insight into the usability of a product. Nielsen makes the point that after the fifth participant, most of the problems have already been discovered and any subsequent participants will only repeat and confirm what is already known. This is echoed by Dumas & Redish (1993:128) who say that "after you've seen several people make the same mistake, you don't need to see it a 10th or 20th or 50th time". As we are comparing two versions of an interface (i.e. one user guide with iconic linkage and one without) we will, of course, need twice the number of participants. This fits in with claims by Dumas & Redish that usability tests generally involve 6-12 participants divided among a number of subgroups. A problem with this, however, is that using a small number of participants renders it very difficult to perform statistical analyses to establish the statistical reliability of the data collected. Many of the more commonly used statistical tests such as t-tests etc. require much larger numbers. There is, however, a way around this using the Wilcoxon-Mann-Whitney exact test. This test will be discussed in more detail later.

Usability Evaluation Procedures

In the previous sections we discussed the various approaches, methods and techniques used in usability evaluation. This section examines how these various factors are implemented as part of a usability evaluation. The purpose of this section is to discuss how and why these tools are used in an evaluation. The procedures for conducting evaluations will also be discussed with reference to a number of case studies which relate directly to our examination of user guides.

Practical Usability Evaluation

While having a firm understanding of the methods and tools described in previous sections is, without doubt, essential in conducting usability evaluations, the success or failure of an evaluation depends on a variety of preparations being carried out. Without these preparations, the test will stumble awkwardly on, with no clear goal, purpose, transparency or logic. In the following sections, we discuss the various practical tasks that form a central part of usability evaluation.

Measuring Usability

As we discussed in Chapter 4, usability involves not only how well users can use an interface, but also how users perceive the interface and what their subjective opinions of the interface are (Dumas & Redish 1993:184). We also know that the usability of an interface can be measured in absolute terms on the basis of predefined performance criteria or it can be compared against that of another interface. However, the term performance is vast and can refer to a multitude of elements and factors of interface usage and the execution of tasks. Unless we define which aspects of performance we want to measure, it will be difficult, if not impossible, to determine the usability of an interface.

Wixon & Wilson (1997:664) present a list of what they term "usability attributes" which are characteristics of an interface. These attributes can be used to categorise and quantify the various facets of an interface's performance. The attributes proposed by Wixon & Wilson include:

- usefulness
- efficiency
- memorability
- advanced feature usage
- flexibility
- learnability
- error rates
- first impressions
- satisfaction or likeability
- evolvability

It is apparent that there are some questionable inclusions in this list such as "usefulness" which, as we have already discovered in Chapter 4, is quite separate from usability in that it is a social or commercial factor which does not effect how well users can use something.

On the basis of these categories it is possible to compile a list of measurement criteria for use in usability evaluations. These criteria make it possible to count or measure individual behaviours on the part of the user. Dumas & Redish (1993:184) point out that counting instances of user behaviour requires careful attention but, importantly, that it does not require judgmental decisions; either the event occurred or it did not provided, of course, such events are observable and discrete. Wixon & Wilson (1997:666) provide the following list of measurement criteria:

- time to complete a task
- number of subtasks completed
- time needed to complete a task after a specified period of time away from the system
- number of steps required to complete a task
- number of times users access documentation or technical support
- number of tasks completed
- number of errors per unit of time
- time spent recovering from errors versus time spent working productively
- number of negative reactions to interface
- number of commands or icons remembered after task completion

Dumas & Redish (1993:185) provide their own list of criteria which, although similar in certain respects to the criteria of Wixon & Wilson, are more detailed in their formulation. These criteria include:

- time to complete a task
- time spent in online help
- time spent reading manual
- the number of incorrect menu choices
- the number of incorrect icon choices
- the number of other errors
- the number of calls to technical
- time spent navigating menus
- time spent finding information in manual
- time spent recovering from errors
- the number of incorrect choices in dialog boxes
- the number of incorrect function key choices
- the number of repeated errors
- the number of help screens

- support or for assistance
- the number of repeated visits to
- the same help screens
- the number of times a quick
- reference card is consulted
- the number of searches in the index on each visit to the manual
- observations of frustration
- observations of satisfaction
- viewed
- the number of times a manual is
- consulted
- the number of pages looked at on
- each visit to the manual
- the number of searches in the table of contents on each visit to the manual
- observations of confusion

It is clear from these lists, not to mention other similar lists such as the one compiled by Preece (1994:405), that there is a wide range of criteria which can be measured. Obviously, recording data for all of these criteria would be extremely time-consuming, if not overwhelming. Indeed, Wixon & Wilson (1997:667) maintain that the number of criteria should not overwhelm the test team. They go on to say that 2-3 criteria are sufficient to measure usability. In contrast, Dumas & Redish (1993:185), while acknowledging the impracticality of using all of the criteria, do not restrict the number like Wixon & Wilson. Instead, they say that not all of the criteria are applicable to each test and that only those that relate directly to the product should be used.

If we refer back to Chapter 4 where we detailed the ways in which Iconic Linkage can improve usability, we can see that the other main attributes of interest are: learnability, retention of information over time, comprehensibility, accessibility of information and speed of processing. In the following section we will look at previous studies to see how they used the procedures, methods and criteria we have just discussed.

Usability Evaluation Case Studies

There are numerous published accounts of usability studies conducted in a variety of contexts. These studies are of varying relevance to user guides but many of them provide valuable insights into the practicalities of conducting usability evaluations.

One such study was carried out by Simpson (1990) and despite the fact that it is a developmental evaluation model and it groups print documentation

together, it does discuss two examples of usability testing which give some useful practical guidance for conducting evaluations. One of the studies involved testing online help while the other involved a computer-based tutorial (CBT).

A crucial question investigators must ask themselves, Simpson asserts (1990:42), is what specific data is sought. Simpson maintains that the deciding factor in choosing an evaluation method is the type of usability information needed. He proposes the following stages for any form of testing (1990:45):

- define the test question
- decide what data is needed in order to answer these questions
- select methods for getting this data
- plan how the methods should be implemented

By his own admission, this process is rarely as straightforward as it seems. Beyond this overview, however, Simpson provides little useful practical advice.

Another study, carried out by Harrison and Mancey (1998), compares two versions of an online, web-based manual and examines the optimum elapsed time before gathering users' reactions to the different designs. Rather than examining textual or content-related factors, this study compared different navigation models and their effect on usability. Although this study also treats online and print documentation identically and its objectives are dissimilar to our objectives here, it provides a useful insight into procedures for gathering data using user surveys.

As a way of testing how well users learn and remember information from a manual, the study used a series of questions based on the information contained in the manual. There were eight groups of twelve questions which took the form of cloze tests which could be answered with a single one, two or three word response. Such a method could be used to test the notion that usable texts promote the retention of information over time.

Interestingly, this study also utilised a written script for researchers to follow during tests to ensure consistency for all subjects. The authors do not, however, give any details of the actual tasks involved or the efficiency and error criteria employed (if any). This can be attributed to the fact that the aim of the study was not actually concerned with measuring usability *per se*.

The main finding of the study was that the length of time a user spends working with a product before being asked to give an evaluation affects the final evaluation. However, the authors found that evaluations stabilised after working with the product for 15 minutes. This also shows that think-aloud protocols, argued to be more accurate because of the immediacy of responses, are unnecessary for the purposes of gauging user satisfaction and opinions as there is no pressing need for immediate feedback.

Teague *et al.* (2001) conducted a series of tests at Intel Corp. in Oregon with the similar aim of establishing whether there were significant differences when users are asked to rate ease of use and satisfaction during and after tests. A total of 28 subjects were recruited to perform a variety of tasks using a range of commercial websites. Tested individually, subjects in the two groups were asked questions at either 30 or 120 second intervals while performing the tasks. The questions were based on seven-point Likert scales and subjects had to answer each question orally during the task. After the task, the subjects were asked to answer the questions again in writing. A third group, who did not answer questions during the task, only answered the questions in writing.

The results of this study appeared to indicate that post-task responses were "inflated" and that users gave more honest and representative answers during the task. Not only is this finding in conflict with Harrison & Mancey (1998), but it can be argued that there were other psychological factors at work which can account for this phenomenon. According to various social psychologists, most notably Asch (1956) and Sherif (1937), conformity and the desire to conform and be accepted can frequently cause people to give "false" or less than truthful answers, even though they do not reflect what a person actually thinks. This desire to conform is most pronounced when subjects are asked to publicly verbalise their responses. In comparison, the need to conform is less obvious where subjects are asked to write down their responses in private (Deutsch & Gerard 1955). Thus, it is reasonable to assume that the "inflated" results in the post-task survey are actually more indicative of the subjects' real ratings than the verbal, concurrent ratings. In any case, it can also be argued that subjects' responses only stabilised after completing the tasks (as mentioned previously by Harrison & Mancey 1998). It is possible that, for whatever reason, the subjects were (unwittingly) biased into giving negative answers because they thought that that was what was expected of them.

Another possible explanation can be deduced from the finding that subjects who only answered questions in the post-task evaluation performed their tasks more quickly than the concurrent groups. The concurrent groups took on average 15% longer to perform the tasks and found the tasks

significantly less enjoyable. We can attribute this to the regular interruption and distraction caused by the questions and the subsequent need to refocus on the task at hand. Such activities require additional cognitive effort and as such increase the workload, fatigue and stress for subjects. It is clear, therefore, that post-task evaluation appears to be a more considerate and indeed accurate means of data collection than any concurrent form of questioning.

In contrast to the preceding studies, Zirinsky (1987) provides a detailed and useful discussion of usability evaluation aimed specifically at printed documentation. Zirinsky starts off by stating that in a usability test involving users, we want users to tell us what they dislike about a product, not what they like (1987:62). The role of testers, he continues is to find problems, not to impress researchers with expert performance. A similar point is made by Redish and Dumas (1993:276) who emphasise that users should realise that they are not being tested.

Zirinsky provides a number of recommendations for those preparing to conduct a usability test. The first of these is that all of the test materials (1987:62) should be edited. As part of the editing process, it is essential that there are no typographical errors, style inconsistencies, grammatical or punctuation errors which can distract subjects or even cause them to doubt the validity of the technical material presented. This leads on to checking the document for both technical content and linguistic accuracy. Zirinsky maintains that a manual will improve by no more than 20% as a result of a review, so the better the quality of the product to start with, the better it will be after being reviewed and tested.

As regards actually conducting the test, Zirinsky asserts that users should remain objective and should be fully briefed about the product and their role in the test. They should only be provided with enough information to ensure that they fully understand what is expected of them. Subjects should not be told what the researchers are looking for, i.e. they should be told that they are looking to see which of two versions of a user guide is better, not that we are looking to see what effect repetition has on a document's usability. Furthermore, subjects must be made to feel relaxed and confident enough to make constructive criticisms and comments regarding the document.

It is clear from the previous studies that many of the approaches simply do not apply completely to this study, even though several of the studies provide useful practical pointers. Of the literature reviewed, only two studies specifically set out to conduct comparative usability tests on print documentation where the object is to gauge the effect of a single, non-technical variable. As such, these studies provide a broad framework or model for

conducting an empirical study to test the effect of Iconic Linkage. The first of these, conducted by Foss *et al.* in 1981, aimed to improve usability and accelerate learning by examining the use of supplementary information and the effect of restructuring a user guide. The second study was conducted by Sullivan and Chapanis in 1983 and was concerned specifically with reengineering a user guide to take into account best practice in terms of technical communication and human factors. The following sections describe these studies in detail.

Foss et al. 1981

Basing their work on the claim that previous work in the area of text accessibility and usability was vague and often contradictory, the authors set out to "understand better the acquisition, representation, and utilization of knowledge by novice or occasional users" of software and to "test some ideas derived from current views of memory and attention" (Foss *et al.* 1981:332).

The basic hypothesis is that users learn more effectively when they understand what they are doing. To test this, a comparative experiment was carried out using two versions of a computer manual: both versions were essentially identical in content but one was restructured to present information in order of progressing complexity. Two groups were formed and each was given one of the two versions of the manual.

In addition to the original and revised manuals, the authors produced what they call an "*Advance Organiser*". This document consisted of an explanation and description of the basic characteristics of the software being used. It did this in general terms describing key parts of the software without referring to specific elements of the software. The *Advance Organiser* was given to half of the subjects in each of the two groups.

In conducting the experiment, which took three hours per subject, subjects were first given general information about the experiment and their typing speeds were measured. Selected subjects were given the *Advance Organiser* and told to study it. Subjects were then given one of the two manuals and told to study it for 15-30 minutes. Once this had been completed, a booklet containing nine text editing exercises was distributed to subjects.

The subjects were told to rely as much as possible on the manuals and that they could only ask the experimenter as a last resort. During the course of the experiment, the experimenter remained in the room noting certain aspects of the subjects' behaviours such as the amount of time they spent performing each task, the number of interactions between the subject and

the experimenter as well as the number of tasks completed within the allotted time. Although subjects were told not to ask the experimenter for help, some did in fact ask for help. Such instances were dealt with according to a strict procedure. Firstly, subjects were told that the information they required was contained in the manual. If this was unsuccessful, the experimenter pointed out the specific chapter in the manual. If the subject still experienced difficulties, the experimenter gave explicit instructions.

Such an approach may seem inappropriate when the purpose of the experiment is to determine the performance of subjects using the manual. However, if we look at this in a wider context we can see that if a manual is effective, there should be no need for subjects to ask questions. Thus, a subject being forced to ask certain types of question indicates some form of problem in the manual. Sometimes during the experiment, the experimenter intervened when it was apparent that the subject was struggling to complete tasks, e.g. the subject sat for long periods without accomplishing anything.

In addition to the manual recording of subjects' behaviours by the experimenter, the experiment also utilised a rudimentary (by today's standards at least) keystroke logging application which recorded the type and number of commands used to perform each task.

On the basis of the data collection techniques employed in the experiment, it was possible to analyse performance on the basis of the following criteria:

- number of tasks completed
- average time to complete a task
- average number of commands used to complete a task
- average number of verbal interactions during a task
- number of errors at the point where the subject said the task was complete

The results of the experiment showed that the organisation of a manual can dramatically improve user performance. It also showed that providing advance information allowed users to "scaffold"[6] their learning, thereby making learning more efficient. While this study does not relate directly to the

[6] For a more detailed discussion of scaffolding in the context of social constructivist theories of learning, see Kiraly (2000).

current research in that it involves restructuring the sequence of information in a manual and the use of external materials, it does provide a useful insight into usability evaluation procedures involving documentation.

Another case study which relates more closely to this research in terms of goals and materials is the one carried out by Sullivan and Chapanis in 1983.

Sullivan & Chapanis 1983

Like Foss *et al.* (1981), Sullivan and Chapanis set out to investigate ways of improving the usability of computer documentation by evaluating different versions of a user guide for a text editing software application. However, this particular study differs from that of *Foss et al.* on a number of points: firstly, this study involves the creation of an entirely new manual from scratch; secondly, the study is concerned with the content and formulation of information in the manual as opposed to the organisation of information or the use of supplementary sources of information.

According to the authors, the purpose of this study was…

> …to improve an existing computer manual through the application of rules of document preparation, to measure the extent of the improvement and, based on that experience, to propose a general methodology for the preparation of documentation (Sullivan and Chapanis 1983:114).

The study consists of two broad components. The first component involved producing an "improved" manual on the basis of a review of literature on writing instructions. This involved implementing style guidelines such as using short, active sentences, simple language, sequencing events in the order they are performed etc. The second stage involved conducting a comparative analysis of the original manual and the new version of the manual. Both of these stages involved an element of experimental evaluation and are described in greater detail below.

The experiment was carried out over two sessions for each subject. In the first session, subjects familiarised themselves with the computer using a keyboard learning program. In the second session, subjects were randomly chosen to use either the old manual or the new manual. The first task was allocated 2.5 hours but subjects were informed that they should not feel pressurised because most people would be unable to complete all of the tasks within the allotted time. When this task was completed, subjects were given a

15 minute break before commencing the second task which lasted for an hour.

During the tasks, the subjects were told that if they encountered a problem, they should first try to solve it themselves. They could only ask the experimenter if they really could not solve a problem themselves. If subjects asked general questions, the experiment referred to the manual. If subjects asked specific questions, the experimenter gave specific answers.

When the tests were completed and all data had been gathered, performance was measured on the basis of the following criteria (Sullivan & Chapanis 1983:119):

- quality of text produced
- number of different commands used successfully
- number of questions asked
- type of questions asked

Recognising the importance of subjective user attitudes to overall usability levels, the authors used a post-test questionnaire to determine subjects' attitudes to the manual, the software and the tasks. However, attitudes to the software and tasks were not related to which manual was used (*ibid.*:122). In contrast to the initial questionnaire, the one used here consisted of thirteen 7-point Likert scale questions. Four of these questions related to the manual while the remainder related to the software and the tasks.

The authors found that the subjects who used the new manual performed significantly better in the tasks than the subjects who had used the old manual and that the new manual group were able to use more commands successfully without the assistance of the experimenter. They also found that the group using the old manual asked more than four times as many questions as the other group and that this indicated serious problems with the old manual.

The questions asked by subjects during the tasks were categorised as follows:

1. manual organisation problems
2. global misunderstanding
3. not knowing what command to use
4. failing to distinguish between writing and editing modes

5. not knowing how to use a command or code
6. system-dependent problems
7. task clarification questions
8. negative comments

The results of the test showed that in all categories except category 4, the subjects who used the new manual asked far fewer questions than the other group. In category 4, the number of questions remained constant.

The overall findings of the study showed that it is possible to improve the usability of user documentation through a combination of "local" production rules and iterative production processes. The authors also maintain that the involvement of potential users in the development process is essential.

Experiment to Test the Impact of Iconic Linkage

Having looked at a range of assessment methods and case studies the next step is to put this knowledge into practice to evaluate the effect of Iconic Linkage on the usability of a user guide. The following sections describe the preparations, rationale and methods used to assess the effect of IL on a user guide by simulating the text production phase of the translation process.

Pilot Study

A pilot study is an invaluable part of any empirical study as it not only gives an indication of the likely results, it, more importantly, provides an opportunity to test the methods, materials and processes. Pilot studies are vital in identifying potential problems with the study and make it possible to refine methods and so avoid costly (and embarrassing) mistakes in the main study. Although Iconic Linkage can be implemented in both monolingual text production situations and in translation situations, the study was restricted to examining Iconic Linkage in a single language environment to eliminate the potentially confounding influence of an additional language and the translation process. In addition to reflecting the production of user guides in a single language, this also reflects the introduction of Iconic Linkage during

the editing, proofing and correction stages of both original language documents and translations.

For the purposes of this study, usability (or more specifically, factors which indicate usability problems) was measured on the basis of the following quantifiable criteria (drawing on those discussed on page 197):

• Subtasks completed	• Tasks Completed
• Times user guide used	• Negative reactions to system
• Incorrect icon choices	• Commands and icons remembered afterwards
• Verbal interactions/questions during tasks	• Incorrect menu choices
• Searches in TOC each time user guide used	• Searches in the index each time user guide used
• Observations of confusion	• Observations of frustration
• Errors at point where subject thinks task have been completed	• Observations of satisfaction

Table 2: Pilot Study Usability Criteria

It was also decided to record the times for each of the following criteria:

• Completion of tasks	• Recovering from errors
• Navigating menus	• Using online help

In addition to these objectively quantifiable criteria, a subjective user satisfaction questionnaire and a post-test survey were also administered to examine usability in terms of user satisfaction and retention of information over time.

The pilot study was carried out in four major stages:

1. Preparations
2. Familiarisation
3. Testing
4. Post Test Surveys

Preparations

Preparations for the study involved the creation of a wide range of documents, forms and other materials including the two versions of the user guide. This stage also involved profiling typical users and recruiting participants.

Preparing the User Guides

The single most time-consuming and labour-intensive part of the preparations involved sourcing and editing the user guide. Firstly, a suitable user guide had to be found. The main challenge here was finding software that was both suitable for use in the study and which had documentation with the appropriate content and textual features (i.e. low levels of latent Iconic Linkage). The expression "appropriate content and textual features" encompasses a number of different factors. Obviously, the user guide needed to be of an acceptable professional standard from a quality perspective. Quality was assessed on the basis of completeness of information, spelling, readability, layout etc.

Secondly, and perhaps most importantly, Iconic Linkage is only of use where the same or similar information is presented at various points in the text. The challenge here was to find a user guide where the instances of non-isomorphic, but semantically identical text were not so closely packed together as to render the learning aspect of Iconic Linkage irrelevant (e.g. identical information in consecutive sentences) yet not so far apart as to necessitate reading hundreds of pages of text.

The nature of the software itself plays a crucial role in the presence of these features in that it is simply not possible to create full Iconic Linkage in the text where there is no similarity in the functions or procedures in the software.

Selected Test Software

The user guide selected for use in the study was an extract from the documentation for the *DigiTake* parliamentary recording suite developed by *DigiTake Software Systems Ltd.* In addition to allowing the use of their software, the company also provided invaluable technical assistance in using the product as part of this study. DigiTake is a suite of digital audio recording applications used to record the proceedings of meetings, debates etc. in digital audio format. This software is designed for use in venues such

as parliamentary chambers, large meetings, tribunals etc. and is currently in use in the Irish Parliament (Dáil Éireann).

The extract used deals with an application called DigiLog. This application is used in conjunction with the *DigiTake* digital recording package and is used by parliamentary reporters to log the proceedings of meetings etc. As each speaker rises and begins speaking, the reporter types the speaker's name either manually or using a shortcut and then types the first 4-5 words spoken. Once this is done, the reporter waits until the next speaker takes the floor and then repeats the process. This is repeated for each speaker. In effect, the reporter logs the turn-taking in a meeting to produce a log which serves as an outline when preparing the full transcript of the proceedings at a later stage. The study emulated this procedure with the participants playing the role of the parliamentary reporter.

This software was regarded as particularly appropriate because, although it is based almost exclusively on industry-standard technology and uses word-processing technology familiar to most users, it is designed for a very specific area of application. Consequently the procedures, rather than the technology, are quite specific and specialised and are unlikely to be familiar to participants. This is important in order to ensure that participants are relying on the user guide and not on their own prior knowledge or problem-solving skills. The familiarity of the word-processing environment serves to provide reassurance for participants and also to eliminate the need for more detailed profiling of participants' previous knowledge. Another advantage of *DigiLog* is that various steps and procedures are common to different processes. This means that the likelihood of similar information occurring in the user guide is high and as such, the user guide is more likely to benefit from iconic linkage.

Versions of User Guides

Once the software and user guide were selected, the first step was to produce two versions of the user guide: one with Iconic Linkage and the other without Iconic Linkage.

The first stage of this process was to create a copy of the original user guide and ensure that both versions were identical in terms of fonts, layout and graphics. The format of the documents was updated to make the information even clearer. Given the fact that only an extract of the complete user guide was going to be used and that the available time was limited, additional graphics were included to provide greater clarification for certain sections of the text where otherwise users would refer to the system or another section of the user guide.

The next stage involved rewriting one version of the user guide to introduce Iconic Linkage. In practice, there are a number of ways of introducing Iconic Linkage, including the use of *style guides* and *text processing software*.

Style Guides

As described in Chapter 4, style guides are a good way of implementing usability principles. Style guides also represent an effective way of restricting the ways in which information can be formulated; much in the same way as controlled language does (Power *et al.* 2003:115). In effect, this facilitates the introduction of Iconic Linkage by creating an environment where isomorphic formulations are preferred. So, to implement Iconic Linkage we would need to develop a reasonably detailed and comprehensive style guide to specify the ways in which information can be formulated. There are, however, limitations in the use of style guides. It may not always be possible to develop a style guide because this is itself a lengthy and time-consuming process. As Dumas & Redish maintain, any of the publicly or commercially available style guides can be used to eliminate this problem (1993:60).

Another potential problem with style guides is that if Iconic Linkage is to be implemented on a large scale in a text, a large number of rules will be needed. This introduces the problem of learning, memorising, practising and consistently using the rules. Depending on the number of rules, it may be impossible for one person to remember them all and then to use them consistently without some form of tool to ensure adherence to the style rules (see Schwitter *et al.* 2003). Ultimately then, the best course of action is, perhaps, to opt for a fairly limited style guide which is easy to remember and implement.

Text Processing Software

Text processing software as used here refers to translation memory (TM) tools such as *Trados Translator's Workbench*, *STAR Transit* or *Atril Déjà vu*. However, it is conceivable that some other form of text storage and recognition software or text database could be used (Buchanan 1992). For the purposes of this study, Trados Translator's Workbench was used.

Under normal circumstances, TM tools are used to help translators translate from one language to another, but they can also be used to translate intralingually or from one variety of a language to another, e.g. US to UK English. When rewriting a text using Trados, we have the source text which we will call ST and we have the edited text which we will call the *target text* or TT. As each sentence is edited and rewritten, the TT is stored

along with the ST in a database. Each subsequent sentence in the text being edited is compared against the ST/TT units in the database to see if there is a match. As more and more of the text is edited, more ST/TT units are added to the database and the likelihood increases that a new sentence will at least partly match one or more of the ST/TT units already in the database.

Relating this to Iconic Linkage, while Iconic Linkage by definition refers to making isomorphic, sentences which are non-isomorphic and which cannot be detected by TM tools because they analyse surface structure and not meaning, TM tools can detect instances of partial matches and isomorphic sentences. That TM tools can detect partial Iconic Linkage is clear but what makes this useful is that very often, these partial matches can sometimes be turned into full Iconic Linkage by rewriting them.

Also, TM tools can indirectly detect non-isomorphic but semantically identical sentences thanks to the incidence of placeables. Placeables are words or phrases such as product names, toolbar icons, menu options and dialog boxes etc. that do not change in either editing or translation. As such, they will stay the same regardless of the way in which a sentence is phrased. Thus, if an existing unit contains the placeables X, Y and Z, a new sentence that has these terms may, depending on the ratio of these terms to the total number of words in the sentence, be offered as a partial match solely on the basis of these placeables. From preliminary tests using this method, fuzzy (partial) matches above approximately 60% can frequently represent instances of non-isomorphic semantic matches. They can be rewritten to introduce full or partial Iconic Linkage where there was none before. Trados also provides a concordance search function which allows users to search the database for instances of a particular word or phrase. This function can also be used to identify potential candidates for partial Iconic Linkage.

The primary benefit of this method is that TM tools can "remember" and analyse a greater number of sentences than a human could ever hope to do. In doing so, TM tools capitalise on latent Iconic Linkage, be it full or partial. However, it should be noted that this method on its own can only detect latent Iconic Linkage in the text. This can then be reused throughout the text or transformed into full Iconic Linkage depending on the human operator's memory. TM tools alone cannot introduce new Iconic Linkage into a text, only repeat existing formulations consistently throughout the text.

When style guides and TM technology are used together, however, they form a powerful suite of methods with each one effectively cancelling out the shortcomings of the other. Thus, style guides can be used to introduce new Iconic Linkage into a text by specifying how something should be phrased while TM tools ease the burden of analysing and remembering large amounts of text.

In this study, one version of the user guide was rewritten using Trados in conjunction with selected style guide rules. As already stated, it is not feasible to develop a comprehensive style guide specifically for this study so it was decided to use a commercially available style guide published by *Microsoft Press*. The *Microsoft Manual of Style for Technical Publications Version 3.0* was chosen because of its comprehensiveness and because the ubiquity of Microsoft products makes its writing style more familiar to users. It would be unrealistic and impractical to implement every rule contained in the style guide. Indeed many of the rules contained in the style guide simply did not apply to this user guide. Rather, the following series of rules were selected on the basis of their applicability to DigiLog and how easily they could be implemented:

- parallel structures in headings, tables lists, sentences and procedures
- use gerunds in headings and not infinitive verb forms
- avoid anthropomorphism
- avoid use of Latinisms and foreign words and phrases
- short, uncomplicated sentences
- logical procedural syntax
- use the active voice
- use second person to directly address user
- non-sexist language / bias-free pronouns
- do not use possessive form of product names
- use present tense
- use positive constructions where possible
- use consistent terminology and avoid polysemy

Rewriting was carried out using the two methods in parallel. Thus, as each sentence was analysed in Trados, various style rules were implemented where appropriate and the edited sentence added to the database.

Once the rewritten version of the user guide was completed, both versions were carefully proofed to ensure correct spellings, consistency of formatting, completeness of information, correct references and terminology as well as general style and register. To ensure that both user guides were adequate, fully functioning documents from a technical communication point of view. Flesch readability tests were carried out on each text using the *Microsoft Word* readability statistics function. The results of these tests showed that both user guides were of virtually the same level of readability and that either one would have been perfectly acceptable for users (see Section 2). Figure 1 and Figure 2 show the results of the Flesch readability tests. According to Microsoft Corporation (Microsoft Word Help 2002), the ideal score for documents is between 60 and 70 for the Flesch Reading Ease test and between 7.0 and 8.0 for the Flesch-Kincaid test.

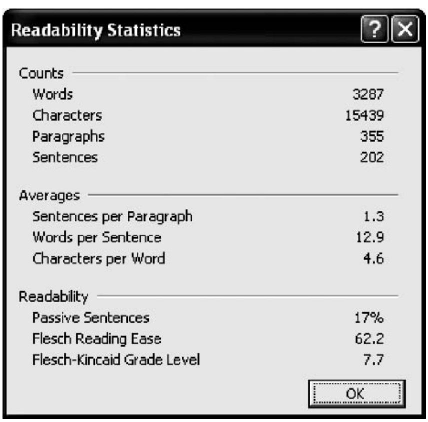

Figure 1: Readability Statistics for Original User Guide

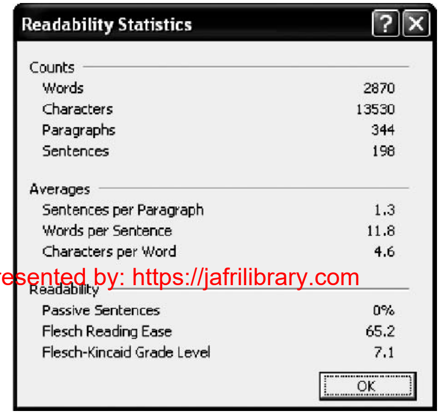

Figure 2: Readability Statistics for Edited User Guide

The readability data also shows that, in addition to reducing the overall word count by approximately 12.7% from 3,287 words to 2,870 words, there are no passive sentences in the rewritten user guide. This is an obvious result of the rewriting process when we consider that one of the rules selected to implement Iconic Linkage explicitly states that the active voice should be used instead of the passive voice.

Nevertheless, it is possible that, even before we conduct the experiment, the elimination of passive sentences could be regarded as a confounding factor. This may or may not be a valid proposition. One argument to support this may be that eliminating passive sentences merely makes the text more readable, not usable. However, referring back to the discussion of readability in Chapter 2, it is apparent that readability is just ***one factor*** which contributes to usability, and as such is not distinct from it. Thus, any improvements in readability (which, in this case are negligible) are as a result of an attempt to improve usability, i.e., the selected strategies aimed at implementing IL.

In any case, if the results of the experiment do show that users using the rewritten version of the users guide perform better, the nature of the improvement in performance will indicate whether eliminating passives is a genuine confounding variable. If, for example, participants in the experimental group only perform better in terms of the speed with which they work, then it is possible that the missing passives are a confounding variable because we can attribute the increased speed with ease of comprehension caused by improved readability or a shorter text. However, if improvements take place across a range of usability criteria, e.g. speed, error rates,

retention of information over time, numbers of errors and satisfaction, then it would be difficult, if not impossible, to attribute such an improvement to improved readability alone. Thus, the issue of whether the elimination of passives can only be answered by the results of the experiment.

These readability scores, while showing that both versions were of a relatively equal standard from a traditional readability point of view, do not show the deeper and more fundamental textual differences between the documents. Instances of full Iconic Linkage can, however, be uncovered using the “Analyse” function in *Trados Translators Workbench*.

Match Type	Original Version	Edited Version
Repetitions - Segments	10	28
Repetitions - Words	21	291
IL Percentage	0	10

Table 3: Comparison of Repetition in Both User Guides

Table 3 contains results from the “Analyse” function run on the original and edited versions of the user guide. While both versions are virtually the same in terms of content and readability, the results clearly show that the amount of repetition or Iconic Linkage in the two versions varies quite significantly. We can see that in comparison to the original version, at least 10% of the total word count of edited version consists of Iconic Linkage as suggested by the fact that 28 out of the 306 segments are repetitions of other segments. (Note: Some partial repetitions could be instances of IL but they are not detected by the analysis tool in Trados).

The finished user guides were proofed one final time before being printed in colour on high-quality paper. The user guides were spiral bound to make them easier to open and place on a desk while reading.

Creating a User Profile

In order to select appropriate participants for the usability study, it was necessary to first identify the type of people who would make up the real users of the software being used. For the purposes of creating this profile, a user profile based on that presented by Dumas & Redish (1993:129-133) was developed (see Appendix 1). The purpose of this was to determine certain characteristics of the users and their backgrounds, skills and qualifications.

This questionnaire was completed by the software manufacturer in consultation with a real user of the software. The questionnaire showed that users are always graduates with a beginner's to intermediate knowledge of computing. None of the users could be defined as experts. The typical user has a background in an area other than politics or computing and all were recruited on the basis of their English language skills. The user profile questionnaire was subsequently used as the basis for selecting participants for the study later on. This will be described below.

Creating an Informed Consent Form

To ensure that all participants in the usability study understood their role in the evaluation process and to ensure that informed consent was obtained, it was necessary to produce a detailed consent form (see Appendix 2).

This form sets out to explain the nature of subjects' participation in the study and explains what type of data is recorded and how it will be used. To guarantee subjects' welfare and to ensure that they do not feel pressurised, stressed or otherwise influenced, the consent form sets out subjects' rights and entitlements.

Safeguarding Tester Identities

Each participant or tester was assigned a unique identifier known as the *Tester ID*. All documents, files, tapes etc. produced in conjunction with a participant's session were marked with this ID exclusively. In order to keep a record of the participants' details and their IDs, a separate Tester ID sheet was maintained (see Appendix 3).

Task Sheets

When conducting a usability test under the time constraints of a laboratory experiment, the participants must be told what work they should perform using the software if the experiment is to be effective. Of course, it is not uncommon for software companies to release advanced prototypes of applications (known as *beta versions*) to a group of users who are simply asked to find errors or problems (Schneiderman 1998:131). However, such an approach, apart from being extremely difficult to standardise, would be impossible to use within the time constraints of a laboratory-based study. Also, open-type "bug hunts" such as this are generally performed by advanced users with more expertise than the participants in our study.

Instead, participants need to be guided through those tasks which will highlight potential usability problems. According to Dumas & Redish (1993:160-163), there are a number of ways of selecting tasks for a usability study. The first way is to use tasks suggested by designers. The authors argue that because designers know the system intimately, they will have a good idea of where potential problems lie. However, it is precisely this detailed knowledge that can result in designers focussing on more complex and advanced problems while overlooking the lower level problems likely to affect non-expert users.

The second way, which was adopted here, involves basing tasks on what real users will actually do with the software. So, for instance, users of DigiLog would create a new log, change the autosave settings, format the text, save the logs as well as actually logging a debate or meeting. A crucial factor in choosing tasks is to determine the size and scope. If an individual task is too short or too simple, it may take just seconds to complete and will be very difficult to observe and quantify. On the other hand, if a task is too long or involves more than one system concept or procedure, not only do we run the risk of exhausting participants we may find that it is difficult to quantify because of the difficulty detecting where one task or subtask ends and the next one starts. In choosing the tasks for this study, only tasks which corresponded to a single concept and which were as self-contained as possible were chose.

Here, we need to address the issue of exactly how self-contained or independent tasks should be and whether tasks need to be performed in order. Wixon & Wilson (1997:670) describe two methods of presenting tasks: *results-based* and *process-based*. Results-based tasks require participants to achieve a specific goal without providing any indication of the intermediate stages. Process-based tasks provide participants with details of the various steps involved in completing a task. For the purposes of testing the usability of a user guide we are interested in finding out exactly what happens when users perform tasks, not just whether they complete the task or not. As such, a process-based approach was adopted.

In their discussion of whether process-based tasks should be independent or interdependent, Wixon & Wilson (*ibid.*) say that independent tasks allow participants to move on to the next task regardless of whether the previous task was completed successfully. This contrasts with interdependent tasks where, if participants encounter serious problems, they may not be able to proceed to the next task without some form of intervention or assistance from the test administrator.

In reality, however, the tasks performed by users in a typical working environment are rarely independent, i.e. they are performed as part of a user's strategy for achieving some goal. Frequently, if users cannot complete a task or subtask, they will not be able to proceed or achieve their goals. This presents us with a compromise between convenience in the usability laboratory and realistic tasks which reflect realistic scenarios of use. In this study, it was felt that the nature and length of the tasks would not pose problems which could not be resolved with a minimum amount of interaction with the test administrator, if necessary. Furthermore, the need for realism was deemed to be of greater importance than convenient test administration. For this reason, all tasks were designed to be dependent on the completion of the preceding tasks.

Recorded Speech

Since one of the tasks in the study involved logging a speech, it was necessary to source a recording of a speech. Ordinarily, the DigiLog application would be used in conjunction with its sister application DigiTake to play pre-recorded audio files to which the log files would be appended. However, it was not possible to set up the client-server hardware needed to do this due to space and hardware considerations in the test venue. Instead, it was decided to record a speech and play it during the task independently of DigiLog using another playback device.

As DigiLog is generally used in a political environment, it was felt that the speech to be logged should come from this domain. A range of speeches were examined before one featuring just two speakers – a male and a female – was selected. A speech involving a male and a female speaker was chosen to make it easier for the test participants (who are inexperienced in audio transcription) to differentiate between speakers.

Once the text was selected, it was read out by a male and a female actor and recorded digitally. The recorded conversation was normalised and saved in MP3 format to reduce the file size and saved on CD. The file was played back on a separate PC to the one being used by participants.

Task Event Log

The quantitative data collected from the usability study consists of times and the number of occurrences of certain events. In order to collect this information, an event log consisting of two sections was created. The first section simply provided boxes for recording the times needed to perform each task as well as times for several other activities.

The second section consists of a table listing each measurement criterion. Alongside each criterion, there is a list of numbers from 1 to 30 which is used to record each individual instance of an event. As each instance is observed, another number is crossed out. The total at the end of the test is then recorded. A vertical line is drawn after each task to allow a breakdown of instances for each task. Appendix 4 shows the Task Event Log.

Post-Task User Satisfaction Survey

As we discussed earlier, usability is measured not just on the basis of how well and how quickly users work with software, but also whether or not they like using the software. Subjective user assessment is a way of determining user attitudes to software – just because users **can** use software does not mean that they like it and that the software is considered usable. This is important for a number of reasons. In discretionary use scenarios, i.e. cases where users **choose** to use software, low levels of user satisfaction will prompt users to abandon a particular product in favour of another product. In mandatory use scenarios, i.e. cases where users **must** use the software to do their jobs, low satisfaction leads to absenteeism, high staff turnover and a variety of other complaints from an unhappy workforce.

For the purposes of this study, a modified version of QUIS was used which omitted the questions relating to *system capabilities*, *multimedia*, *voice recognition*, *virtual environments*, *Internet access* and *software installation*. A number of questions from other sections were also deleted because they were not deemed to be applicable to this study. In addition, certain sections and questions were reworded to make them more specific to this study. A number of questions were added in order to explicitly test the hypothesis of the study. Such questions included:

- Language used in the user guide is consistent
- Language used in the user guide is repetitive

These questions sought to determine whether inconsistency or repetition were negative factors which affected the levels of satisfaction among users (i.e., users may become irritated by excessively repetitive text or fatigued/confused by inconsistent formulations).

Post-Test Survey

As discussed early one of the assessment criteria with which usability is measured in this study relates to how well users remember information over

time. Of course, the ideal method for testing retention of information over time is to ask users to use the system again using slightly different tasks which draw on the same information. However, the expense and difficulty in securing participation for an additional test session proved to be prohibitive in this study.

Instead, a variation of the interviewer-administered questionnaire (see page 189) was used. Based on the core knowledge required to perform the test tasks, a series of 15 questions (13 of which were multiple choice) was compiled to test how much users remembered after completing the tasks. The multiple-choice format was chosen to facilitate "cueing" or recognition of answers. While this may seem erroneous in that it "gives" users the answers, it emulates the type of prompting provided by visual clues on-screen in the software such as menu options, icons, dialog boxes etc. Unfortunately, time constraints mean that we cannot determine which answers were "guessed" or triggered by recognition and which ones the testers actually knew without any cueing. Nevertheless, both types of answer combine to show how much information is retained over time. The actual questions used in the post-test survey are presented in Appendix 5.

Test Environment & Tools

The pilot study was conducted in a large and spacious office which was free from clutter, bright and well ventilated so as to provide as pleasant and natural a working environment as possible. Refreshments were provided for participants. The layout of the test equipment and the work area is shown in Figure 3.

The test was conducted on a PC with a Pentium III 2.4GHz processor, 256MB of RAM, a CD-R drive and a 40 GB hard disk. A video camera was positioned to the right of the participant at a distance which allowed the camera to be as inconspicuous and unobtrusive as possible while still being able to record the participant and the surrounding area. The aim of this was to record not just the participant's face but also the user guide while it was being used. This was important in order to record the length of time the user guide was used **and** the way in which it was used (i.e. leafing through the pages or carefully reading each page).

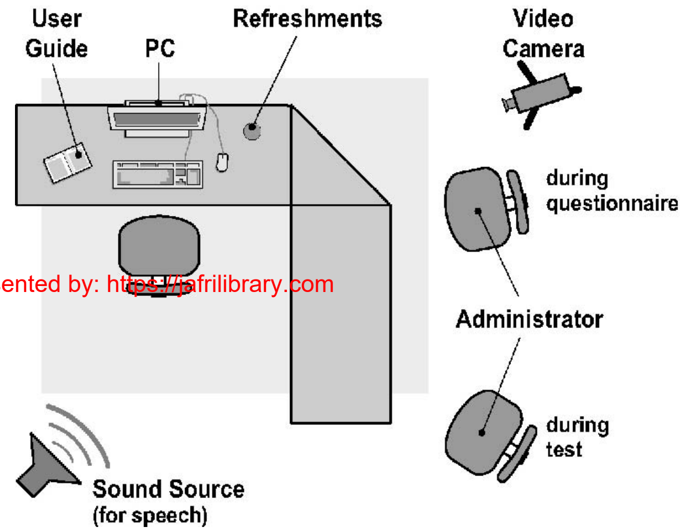

Figure 3: Usability Laboratory Layout

It is also clear from the diagram that the test administrator remained in the room during the test. Although the benefits of indirect observation are clear (see page 186), it was felt that a form of direct observation was necessary for a number of reasons. Firstly, in the absence of a separate observation room with viewing and intercom facilities (see Schneiderman 1998:128) communication with the participant would have been extremely difficult if not impossible. Secondly, having the test administrator in the same room as the participant made it possible for the test administrator to intervene and provide assistance if necessary. This is particularly useful when using interdependent tasks as is the case in this study (as discussed on page 217). It was, therefore, much easier to administer the test.

Two items of software were specially installed on the PC: the test software (DigiLog) and the screen logging software (Camtasia).

Screen Logging Software

While user satisfaction questionnaires and video recording go some way to helping us understand the relationship between users and a system, they only provide part of the overall picture. To fully understand users' performance as they carry out tasks, it is necessary to see what they are doing on-screen, i.e. to observe how they actually interact with the system.

Obviously, it would be ill-advised to sit beside the participants and look over their shoulders as they work, not least because it may affect their performance. It would also be physically impractical and would leave us with no permanent record of events. Some solutions proposed involve positioning a video camera in such a way that it can record events taking place on-screen (Preece 1994:619). Other solutions involve connecting an additional monitor to the computer being used by the participant. Known as a slave monitor, it is usually located in another room and shows what is happening on the participant's screen. This approach can be combined with a video camera to provide a permanent record of events. Neither method is particularly attractive because of the additional hardware and technical requirements.

Another approach is to use some form of software logging tool (Preece 1994:626-627). At their most basic, logging tools can record the keystrokes made by users. More advanced versions also record timestamps for each key that is pressed to give an insight into the time taken to perform tasks. Unfortunately, such products only record actions carried out using the keyboard – they do not record actions carried out using the mouse, e.g. opening menus, selecting items, highlighting text or clicking icons. At the opposite end of the spectrum there are sophisticated tools which take the benefits of interaction logging and video recording.

One such product is *Morae* by Techsmith. This product allows usability researchers to produced synchronised screen and video recordings, observe and mark critical moments and analyze, edit and share recordings. The software can also be used to record audio produced, for example, during think-aloud protocols. As regards user activities, Morae can be used to record changes to web pages, mouse clicks, keyboard input and, using the Remote Viewer, on-screen text and markers. Using such a product would undoubtedly be of great benefit, however, the cost of such a product (around USD$1100 at time of writing) is prohibitively expensive for many people unless they are making a significant investment in usability research.

A more feasible approach is to use stand-alone screen-recording software to record real-time moving images of the events taking place on-screen. Generally, such tools are used to create online training tutorials but they can just as easily be used for the purposes of a usability study. There are several products all of which provide much the same functions such as *My Screen Recorder*, *Matchware*, *Camtasia* and *HyperCam* were evaluated. However, *Camtasia*, which his produced by the same people who make *Morae*, is one of the more suitable options because of its superior functionality, ease of use, image quality and portability.

Camtasia Studio is a suite of tools for recording screen events as well as editing and producing “movies”. These tools make it possible to add annotations, audio files, text etc. to recordings. The suite also includes a proprietary player which is used for showing recordings (the recordings can also be played using Microsoft Media Player although the image quality is much better using the Camtasia player).

Recordings are stored in AVI format and with standard compression levels produce perfect quality files; average file sizes are approximately 1MB per minute of recording. There is also a “pack and go” facility which allows high-quality playback of recordings without the need to install Camtasia. The recording tool can be used to specify the precise area of the screen to be recorded – any section of the screen from a dialog box to a window to the entire visible screen area can be selected. Once Camtasia was installed on the PC, separate folders for storing the recordings were created for each participant.

The benefit of using Camtasia is that it records everything that takes place on-screen – even when text is entered. Camtasia cannot, however, record when function keys or keyboard shortcuts are used – it only records the results of such actions provided they have a visual consequence which is displayed on the screen. However, because DigiLog is a mouse and menu-driven application, this is not a problem. What is more, Camtasia records all actions performed on a PC, not just those carried out in DigiLog. Thus, when users switch between applications or windows or when they use the *Start* menu, Camtasia will record it.

Also, because Camtasia features an elapsed time counter, it is possible to calculate the time spent performing tasks and subtasks, simply by measuring the time between the start of the task (or when the mouse first moves) and the end of the task (when the pointer stops). Each task can be labelled using the effects tool, e.g. a banner can be added to the recording for the duration of a particular task.

Participants

According to Nielsen (2001) and Dumas & Redish (1993:128), a typical usability study will generally feature a maximum of two groups of five people. This presented a problem for the pilot study in that the numbers required are almost the same as for the main study. The minimum realistic number of participants for a group is three to ensure that the results are reasonably reliable and not the result of the idiosyncratic behaviour of one or two participants (Dumas & Redish 1993:128).

With this in mind, two groups of three participants were recruited for the study. Based on the results of the initial user profile (see page 194), it was decided that participants should be graduates, have excellent English language skills and reasonable PC skills. It was not necessary to look beyond the local campus population for participants because such a group was readily available in the form of students on the *Graduate Diploma/MA in Translation Studies* at Dublin City University. In addition to being a postgraduate course, the course also provides thorough training in both general and translation-related computer applications as well as various English language courses.

An email was sent to all students on the programme explaining the nature of the usability study. Students were asked to participate as co-testers in order to help assess the usability of a software product and its user guide. It was emphasised that participants would not be tested, but rather that they were testing a product and giving their opinions on it. As such, there was no need to feel pressurised, concerned or reticent about their performance. Potential candidates were informed that their involvement would require a commitment of three hours over three weeks for which they would be paid €15 per hour.

One stipulation was made in that participants must be native speakers of English. This was necessary to rule out any potential problems caused by various levels of English language skills among the fairly high proportion of foreign students studying on the programme.

Of the nine respondents, six were chosen at random to take part in the study. The others were thanked for their interest and it was explained that the quota had been reached. Each participant was contacted individually and none were made aware of the identity of the other participants. In an effort to minimise the risk of participants discussing the experiments among themselves, the tests were conducted outside term time so that students would not be in as regular contact with each other. In addition, test sessions were scheduled as far apart as possible to rule out participants meeting in or outside the laboratory. Participants were explicitly requested not to discuss the nature of the experiment; this was reiterated in the terms of the Consent Form (see Appendix 2). Test sessions were arranged with each participant and recorded in the Tester ID form.

Session 1: Familiarisation

The pilot study was conducted in three sessions over the course of three weeks for a total of 2.5 hours. Session 1 was the "Familiarisation" stage of

the study and it involved introducing participants to the *DigiTake* package. A product brochure detailing the components of the suite as well as their functions and uses was emailed to each participant. Participants were instructed to spend approximately one hour reading this document. The purpose of this was to familiarise participants with the system so that they would have a context for the information they would learn from the user guide (Foss *et al.* 1981:334). This document aimed to help users understand the software and the tasks they would eventually perform. This document did not, however, provide any specific information on how to use the system or its functions. It simply described the general technology, functions and architecture of the system along with its typical working environment.

Each participant was also contacted to arrange and confirm their test times and dates for Session 2.

Session 2: Testing

Session 2 involved conducting the actual test sessions with participants. Upon arrival in the test laboratory, each participant was welcomed and offered refreshments. Having been shown to the workstation, the purpose and nature of the test was again explained. It was stressed that the purpose of the study was to assess the user guide and not the participants' abilities as regards computers, typing, intelligence etc. Participants were told that they should regard themselves as co-testers – they were part of the test team and that they were doing the testing, not being tested. They were also told that they could take a break or withdraw from the study at any time.

Next, each user was given a consent form (see Appendix 2). They were asked to read this form and ask for clarification, if necessary. If participants were happy with the terms of the consent form, they were asked to sign the consent form. The test administrator also signed the form.

At this point, the video camera was started. Although the tasks had not yet started, turning on the camera at this stage helped participants become accustomed to the camera and the low level of background noise it created before they actually started working. This was intended to minimise any adverse effects caused by the camera's presence.

Once the camera was started, participants were given a randomly preassigned user guide. They were told that they had up to 30 minutes to read the user guide. They were also told that they would be allowed to use the user guide during the tasks. Participants were told that during the test, they should not ask questions or otherwise ask the test administrator for assistance

unless it was absolutely essential. At all times, they must consult the user guide for answers first. If a participant asked a general question, the test administrator referred the participant to the relevant section of the user guide. The test administrator only provided specific information where the question arose as a result of a technical problem with the hardware or software or where, after consulting the user guide the participant was in danger of not completing the task. When participants had finished reading the user guide, a task sheet (see Table 4) was distributed. Participants were informed that the test would proceed one task at a time and that they must not start a new task until told to do so. Upon completion of each task, the participants were asked to inform the test administrator.

Before starting the first task, the test administrator started the Camtasia screen recorder. This simply involved having Camtasia running in the background before the participants arrived and then pressing a single function key on the keyboard to start recording. The first task was explained orally and participants were directed to the task sheet and instructed to commence the task. Upon completion, the participants informed the test administrator and were given the opportunity to ask questions, give comments or take a break. Each task was conducted in this manner.

When all of the tasks had been completed, Camtasia was stopped and participants were again given the opportunity to take a break. Participants were asked whether they had any initial comments on the test. The administrator did not discuss these comments because the only reason for asking this question was to help put participants in an analytical frame of mind whereby they look back on the tasks from the point of view of a tester. This served to prime them for the QUIS questionnaire which was administered next.

In administering the QUIS questionnaire, the test administrator moved to a seat beside the video camera and directly opposite the participant at the desk (see Figure 3). Participants were told that a core part of usability testing was to find out how users feel about using a product. They were told that just because something is easy to use, it is not necessarily usable if users do not like using it. Participants were reminded that there were no right or wrong answers. The participants were given the questionnaire and asked to complete it. If any questions were unclear or confusing, participants were told to ask the test administrator.

Participants were then thanked for their assistance and the importance of their participation was emphasised. At this point, the video camera was turned off. During the first and second sessions, however, the camera was turned off immediately after the QUIS questionnaire was administered. It

soon became apparent, however, that informal comments which sometimes continued right up to the point the participant left the room were not recorded. For this reason, the camera was left on until each remaining participant left the laboratory.

Session 3: Post Test Survey

Session 3 involved administering the post test survey. These sessions were carried out exactly one week after each participant's second session. The post-test survey involved administering the multiple-choice test sheet (see Appendix 5) and took approximately ten minutes. Participants were given the opportunity to ask questions about the study as a whole. Participants were thanked for their participation which was now at an end.

Findings of Pilot Study

The results of the pilot study can be grouped into the following categories:

- Results of Time Measurements
- Error Rates
- Results of QUIS Questionnaire
- Implications for Experimental Design

Results of Time Measurements

In calculating the times for the various activities set out in the Task Event Log, it soon became apparent that this list contained criteria and aspects of subjects' performance which could not be detected easily, if at all. Problems arose mainly because of the difficulty in establishing exactly what a subject was doing at a given point in time. Thus, it was not always possible to distinguish between a subject who was recovering from an error and a user who was performing a task normally.

Similarly, determining when users could be regarded as being unproductive was impossible when we consider that users may be actively thinking about a problem or looking for information on screen while giving the impression of inactivity. One way of combating this would be to implement

some form of think-aloud protocol but for reasons described already, the use of think-aloud protocols would be of limited use here.

Another problem which arose in the study was that of different subjects having different levels of manual dexterity. The result of this was that measuring the time subjects spent navigating menus was not realistic quite simply because it took some subjects longer to physically make their way through the menus. It would seem more appropriate to treat menu navigation as discrete events whereby each occurrence of a subject search through a number of menus apparently looking for the desired option is counted on a "per incident" basis. Ultimately, the only useful and feasible measurement which could be carried out was the time taken to complete each task. Table 4 lists the nature of each individual task.

Task 1	Create two new entries in *QuicKey*
Task 2	Create a new log in *DigiLog* and configure the automatic save settings and set the working directory.
Task 3	Logging task
Task 4	Format the text in the log
Task 5	Manually save the log in RTF format to a specific location

Table 4: Pilot Study Tasks

It was not possible to record times for subtasks because of the difficulty in establishing precisely when one subtask had been fully completed. Some subjects returned to certain parts of tasks which confounded matters. It should also be remembered that Task 3 involved subjects logging a pre-recorded speech. As such, this task always lasted for the same duration (i.e. 5:03 minutes). Consequently, there was little sense in including the times for this task as the fixed time tells us nothing about a subject's performance. This particular task is better used as an opportunity to examine the usability criteria exclusively.

In the case of the Experimental group, problems arose with one subject who failed to follow instructions regarding reading both the familiarisation material sent out as part of Session 1 and with regard to reading the user guide for the specified length of time during Session 2. Consequently, the subject was unable to complete the tasks and as such, no data could be collected for this particular subject.

Error Rates

Of the large list of criteria on page 198, only some proved to be useful for detecting errors during the test. The following tables illustrate only those error criteria which provided at least one observable instance during the test. The remaining criteria have been omitted. In gathering the data, the scores from the post-test survey were used as data for the criterion "Commands & Icons Remembered Afterwards".

In the main study this data is referred to as the "PTS Score". Once the data has been compiled, all of the figures except the PTS Score and the number of tasks completed were added together. The PTS Score and number of tasks completed were subtracted from this total to give an overall error score. The purpose of this score is to facilitate comparing the number of errors made by each subject. The following list shows the new amended error criteria for the main study.

- Tasks Completed
- Times User Guide Used
- PTS Score
- Incorrect Icons
- Incorrect Menus
- Verbal Interactions
- Confusion
- Satisfaction
- Error Score

Again, owing to the difficulties encountered with one participant, it was impossible to gather data for this part of the experiment. Nevertheless,

when calculating average scores for the Experimental group, the participant was treated as a non-scoring member of the group so that averages could be calculated on the basis of equal group numbers.

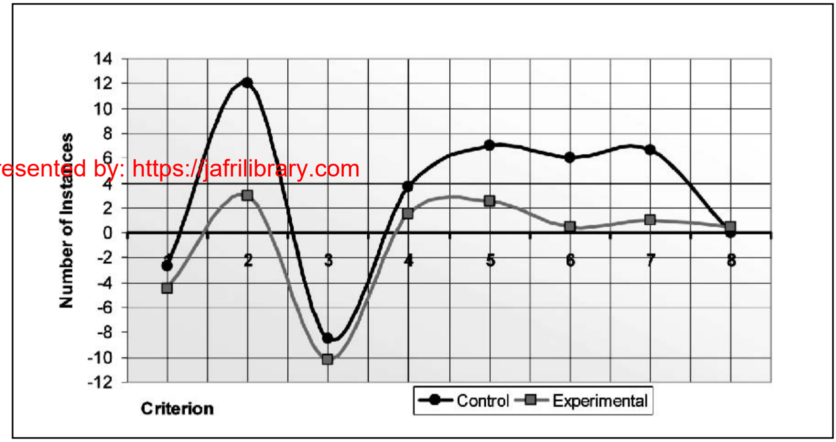

Figure 4: Initial Graph Attempting to Show Pilot Error Scores

It also became apparent when presenting this data in the form of a graph, that there were problems in the way the results were calculated for each criterion. Figure 4 shows that the results for the "Tasks Completed" (Criterion 1) and particularly the PTS Score (Criterion 3) noticeably skew the graph. This fact highlighted faulty logic in the way criteria were worded and applied. The list of usability criteria provided a list of errors which each participant committed. As such, it makes little sense to subtract the number of correct answers given from the total number of errors. Instead, it was decided to define the PTS Score as the number of incorrect answers provided in the post-test survey as this is more compatible with the aim of the remaining criteria. The criterion governing the number of tasks completed was reworded as "the number of tasks ***not*** completed". Figure 5 illustrates the graph with the new data.

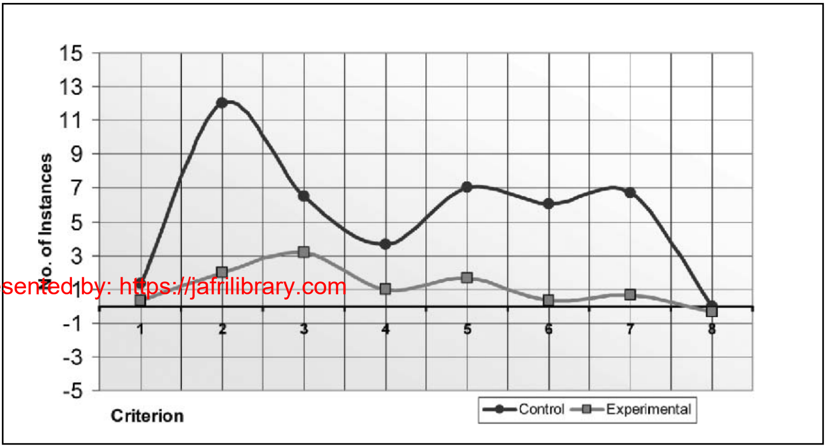

Figure 5: Usability Criteria Error Scores for Pilot Study Groups

Implications for Experimental Design

Over the course of the pilot study and during the subsequent analysis of the data, it became apparent that there were a number of areas relating to the design of the experiment that were either unsuitable, insufficiently refined or which were otherwise problematic. The problems encountered can be grouped under the broad categories of: *Tasks*, *Data Collection* and *Criteria*.

Tasks

There were a number of issues relating to the tasks to be performed as part of the test which gave rise to certain problems. The first problem was in relation to the manner in which subjects were prepared for the tasks. Central to the entire test was the need for all participants to read the user guide for 30 minutes. It emerged during the tests that 30 minutes was too long a period of time. Having noted the average times spent by each subject, a period of 20 minutes emerged as a more realistic period of time.

Another problem with the length of time subjects should spend reading the user guide emerged during the session with subject P4. This participant failed to read the brochure provided as part of Session 1 and despite numerous requests, this subject refused to read the user guide for the required period of time, arguing that this was not how the participant normally used a user guide. The participant was asked a number of times to read the user

guide but became agitated and defensive. Consequently, the participant did not read the guide and so was unable to complete any of the tasks. It is clear from this that it is absolutely essential that participants comply with all instructions and that the there needs to be some way of firmly insisting that participants read the user guide for the full period of time. It is proposed that participants be told quite openly from the outset that they must spend the full time reading the user guide and that this is to ensure consistency and the effectiveness of the experiment.

Problems also arose with regard to Task 5 which involved subjects formatting the text in the log created as part of Task 3. It was clear that all subjects had a good level of knowledge of *Microsoft Word* and that this interfered with the solutions subjects chose as part of Task 5. Although there are two ways of formatting text, i.e. using the tool bar or the menus, each subject without exception used the tool bar which used the same icons as *Microsoft Word*. So, rather than using the information in the user guide, subjects were using their existing knowledge of another application to complete the task. It was decided, therefore, to omit this task as it did not provide any meaningful insight into the usability of the user guide. Nevertheless, the task did show that both groups were well matched in terms of computer skills so it indicated that the method for selecting participants was effective.

Data Collection

During the test sessions, a number of areas became apparent where the method of collecting data could be improved. The first such area involved keeping the video camera running until the subjects had left the laboratory. The reason for this was that some subjects continued giving feedback right up until they left the lab. When the video camera had been switched off immediately after the post-task questionnaire, these comments were lost. However, during the analysis of the questionnaire data, it soon emerged that the comments and feedback elicited from subjects, while very interesting from a software and document design point of view, were not particularly interesting from the point of view of assessing the usability of the user guide. As a result, it was decided that this type of information should not be collected during the main study.

It also emerged that *Camtasia* was not suitable for recording entire test times because it was not possible to see what was going on when a subject appeared to stop working. It was impossible to tell whether the subject was reading the user guide, talking to the administrator or staring at the screen. In addition, some subjects immediately began tackling a task on-screen using

the mouse while others immediately began re-reading the user guide. It was decided that the video recordings would be the better for recording times. Task times were recorded from the time subjects were instructed to start to the time the subject announced completion of the task. Furthermore, although Camtasia files can be used to record partial tasks, it is too difficult to do so in practice because some subjects showed a tendency to repeat subtasks. It was decided that Camtasia would still be used to detect incorrect menu and icon choices etc. in the main study, however.

Several subjects experienced difficulties in hearing and understanding the recorded speech used during Task 3. The speech needed to be either re-recorded or digitally enhanced to make it of an acceptable quality. Problems were also caused by unclear instructions and a misprint in the task sheet and imprecise verbal instructions from the test administrator. Before conducting the main study the task sheet was clarified and updated and verbal instructions to participants were properly scripted. This modification to the experimental method had an additional benefit in that it helped control the amount variability which may inadvertently come as a result of unscripted instructions from the test administrator. For example, on any particular day, the administrator may give more or less information to the participant. Scripting the instructions given by the administrator, coupled with the strictly enforced procedure for dealing with questions (see page 226) helped rule out the possibility of the administrator making *ad hoc* comments which might vary from session to session and consequently biasing or assisting individual participants.

Criteria

From the relatively long list of criteria used to assess usability, it emerged that many were unsuitable for use in this study, either because relevant events did not occur with enough frequency to justify their inclusion or because they were impossible to quantify given the procedures and equipment in use. As a result, some criteria were deleted while others were rephrased slightly to make it easier to quantify them. An example of a rephrased criterion is “Errors at point where subject thinks task completed” which was rephrased as “Number of times subject stops work without completing a task”. It was also necessary to add additional criteria as a result of other phenomena which were observed during the test. The pilot study also made it clear that additional work was needed to establish exactly how the criteria should be applied in the main study.

When analysing the error criteria data, it was apparent that the method for handling data from the Post-Test Survey was not appropriate (see page

231). Instead, the number of **incorrect** answers was added to the totals for the other error criteria. Apart from faulty logic, subtracting the number of correct answers from the total errors meant that any attempts to represent the data in the form of a graph produced seriously skewed results.

In addition, it was felt that the application of error criteria was subjective and, at times, inconsistent. This can be attributed to the fact that a single person was responsible for determining whether an incident actually corresponded to one of the criteria. It was decided, therefore, to clearly define each criteria and what each one involved. The following list presents the modified error criteria and the definition of each criterion which must be observed in order to be recorded.

Criterion 1 *Tasks Not Completed*: The failure of a user to complete a task despite the administrator identifying the precise page in the user guide; ultimately resulting in the administrator giving explicit verbal instructions on how to complete the task.

Criterion 2 *Times User Guide Used*: Each occasion where the participants stops working to read the user guide.

Criterion 3 *PTS Score*: The number of questions answered incorrectly by each participant in the post-task survey.

Criterion 4 *Incorrect Icon Choices*: Where a user clicks an icon which is not associated with the task currently being performed.

Criterion 5 *Incorrect Menu Choices*: Where a user chooses a menu option not associated with the task currently being performed OR where a user scrolls through several menus without choosing an option.

Criterion 6 *Verbal Interactions/Questions*: Each occasion where a participant asks a question relating to the way a task should be performed or whether a task has been completed.

Criterion 7 *Observations of Frustration*: Incidents where a participant expresses frustration verbally or where a participant's body language (e.g. sighing) or facial expressions (e.g. frowning) indicate frustration.

Criterion 8 *Observations of Confusion*: Incidents where a participant expresses confusion verbally or where a participants body language (e.g. head-scratching) or facial expressions indicate frustration.

Criterion 9 *Observations of Satisfaction*: Incidents where a participant expresses confusion verbally or where a participant's body language or facial expressions indicate frustration (e.g. smiling).

Criterion 10 *Incorrect Commands/Input*: Incidents where a participant uses an incorrect shortcut key or types incorrect commands into a field.

Criterion 11 *Stopped Work without Completing Task*: Each instance where a participant mistakenly believes the task to be complete or where a participant gives up.

Main Study

The main study was conducted four weeks after completion of the pilot study. By this stage all of the necessary changes highlighted in the pilot study had been implemented.

The venue for the main study was a usability laboratory set up in offices in the Houses of the Oireachtas (Irish Parliament) in Dublin where the DigiTake product is in use. This was made possible by *DigiTake Software Systems Ltd.*, the company who manufactures the DigiTake package. The company also facilitated the recruitment of participants. Owing to the fact that Parliament was in session at the time of the main study, access to both the laboratory and subjects was restricted with the result that it was possible to conduct only one session per day. The main testing sessions took ten working days to conduct.

In terms of preparations, the majority of the work involved modifying the task sheets to reflect the omission of the text formatting task (Task 4 in the pilot study). The task event logs were also updated to reflect the new usability criteria. The following tasks were retained in the main study:

- *Task 1*: Create two new entries in *QuicKey*
- *Task 2*: Create a new log in *DigiLog* and configure the automatic save settings and set the working directory.
- *Task 3*: Logging task
- *Task 4*: Manually save the log in RTF format to a specific location

During the course of the pilot study it also became apparent that the quality of the recorded speech was not entirely satisfactory. Some subjects found the volume to be too low in places and that the occasionally short gap between speakers did not always allow them to finish typing the previous sentence. Rather than re-record the speech again, the speech file was edited using the *Cakewalk 8.0* digital audio package. The volume levels were normalised to enhance sound quality and gaps of 2–3 seconds were inserted

between each speaker's turn to allow subjects to complete typing the previous sentence. The overall layout of the usability laboratory was kept almost identical to that used in the pilot study although certain physical features of the office geography necessitated minor changes.

Tools

As in the pilot study, the workstation consisted of a high-specification PC onto which *DigiLog* and *Camtasia* had been loaded. Again, labelled directories, forms and video cassettes were readied for each subject. A second PC was used to play the recorded speech.

Participants

For the purposes of the main study, it was felt that greater realism could be achieved by conducting the test in the same environment as real users would use the software. In this case, the software would be used in the Irish Parliament. It was also felt that because of the venue for the tests, it would be possible to gain access to subjects who more closely reflect the real users of the software.

With the assistance of *DigiTake Software Systems Ltd.*, contact was made with Fianna Fáil, the political party in Government at the time of the study, and permission was obtained to recruit the party's personnel assistants as subjects in the study. This particular group is much more suitable for participation in the study because, in addition to having a very high proportion of graduates with appropriate levels of computer knowledge, potential subjects are working in the same environment as parliamentary reporters and as such demonstrate an excellent understanding of politics and current affairs. As such, by drawing subjects from this group, it was possible to achieve a much higher level of realism.

The same selection criteria were used in the main study as were used in the pilot study: subjects must be graduates, must be native speakers of English with excellent communication skills and beginner to intermediate level computer skills (i.e. 1-3 years experience). A total of 10 suitably qualified participants were selected at random. The details of each participant were recorded in the Tester ID form and each person was assigned a unique ID number.

Method

The main study was conducted in three stages over four weeks. During the familiarisation stage, the same product brochure was emailed to participants to explain the nature of the product and the context in which it is used. Participants were asked to read the document carefully.

Each participant was scheduled for a test session over the course of the following two weeks. The procedure for conducting test sessions was almost identical to that in the pilot study. However, the changes highlighted above were incorporated into the procedures. In addition, as previously mentioned, the text formatting task was omitted from the test.

Upon arrival, each participant was welcomed and offered refreshments. The purpose of the study was again explained and participants were asked to read and sign the consent form. Participants were then given one of the user guides to read for 20 minutes. At the same time, the video camera was started. After the participants had finished reading the user guide, they were given a task sheet. It was explained that they should move on to the next task only when instructed to do so. A verbal explanation of the tasks written on the task sheet was also given to ensure clarity and to give participants the opportunity to ask questions.

Participants were told that during the task, they should always try to find answers and solutions to problems from the user guide. Participants were told that they could only ask questions or ask for help from the test administrator as a last resort. The Camtasia screen recorder was started and participants were instructed to begin when ready. During the test, where participants asked questions the answers to which were contained in the user guide, the test administrator directed the participant to the relevant page in the user guide. If, after reading the relevant section of the user guide, a participant still encountered difficulties and appeared to be in danger of not completing the task, the test administrator provided more explicit and detailed information; each of these instances counted as a separate interaction with the administrator.

When participants had completed the tasks, Camtasia was stopped and participants were again given the opportunity to take a break. Participants were asked whether they had any initial comments on the test. The administrator did not discuss these comments because the only reason for asking this question was to help put participants in an analytical frame of mind and prime them for the QUIS questionnaire which was administered next. Arrangements were then made to meet each participant exactly one week

later in order to complete the post-test survey. The post-test survey was completed in the presence of the test administrator and lasted on average 10 minutes.

Results of Usability Evaluation

For clarity of presentation, the data collected during the empirical study is

presented in tabular form in Appendixes 6-8. The following sections contain an analysis and discussion of the results and their implications.

Analysis of Results

Having collated and recorded the data gathered during the main study and observed the differences between the Control and Experimental groups, the next step is to ascertain whether such differences are statistically significant or simply due to chance.

The nature of this study and the fact that only 10 subjects were used, poses a number of problems for any attempts at statistical analysis. Generally, the reliability of data is calculated using methods such as a *Student's t-Test*. However, such tests require at least 20-30 observations for the results to be statistically reliable. This means that the data from the main study cannot be analysed using the more common statistical tests.

This problem can be overcome, however, by using a Rank-Sum test, also known as the Wilcoxon-Mann-Whitney exact test. It is used for comparing two populations to detect "shift alternatives". That is, the two distributions have the same general shape, but one of them is shifted relative to the other by a constant amount under the alternative hypothesis (see below). This test can be used for either continuous or ordinal categorical data. It is one of the most popular nonparametric tests for detecting a shift in location between two populations and is "the only form appropriate for comparing groups of sample size 10 or smaller per group" (Helsel & Hirsch 1993:119).

At the simplest level, the Wilcoxon-Mann-Whitney exact test is used to determine whether one group tends to produce larger observations than a second group. The null hypothesis of this test states that the probability of one group (x) producing a higher value than another group (y) is 50%.

This can be written as:

$$H_{0:}\ \mathrm{Prob}[x>y] = 0.5$$

Helsel & Hirsch (*ibid.*) state that the alternative hypothesis is one of the following:

$H_{1:}\ \mathrm{Prob}[x>y] \neq 0.5$	(two-sided test where x might be higher or lower than y)
$H_{2:}\ \mathrm{Prob}[x>y] > 0.5$	(x is expected to be larger than y)
$H_{3:}\ \mathrm{Prob}[x>y] < 0.5$	(x is expected to be smaller than y)

Due to the limited sample involved in this study (n = 5), the usual asymptotic p-values would not reliable, so exact p-values were calculated throughout using *StatXact 4.01* from Cytel Software Corporation, Cambridge MA.

In performing statistical analyses on the three sets of data, i.e. task times, error rates and user satisfaction levels for each group, it was necessary to decide upon the level of detail to be used. While analysing, for example, data for individual error criteria or times for individual tasks would be interesting, the more detailed analyses we perform the more likely we are to encounter a Type I error (Norman 2003). Type I errors refer to spurious results which can look credible but which are in fact wrong or misleading. This problem is exacerbated by the small sample size used in this study. Consequently, it was decided to analyse the overall data for each subject, i.e. total task times, total error rates and overall satisfaction ratings.

Task Times

Total time taken to complete tasks: A two-sided Wilcoxon-Mann-Whitney exact test of the null hypothesis of no difference between total times for the Control group and Experimental group was carried out. This yielded a *p-value of 0.0079* indicating that there is a statistically significant difference between the times for the Control group and the Experimental group.

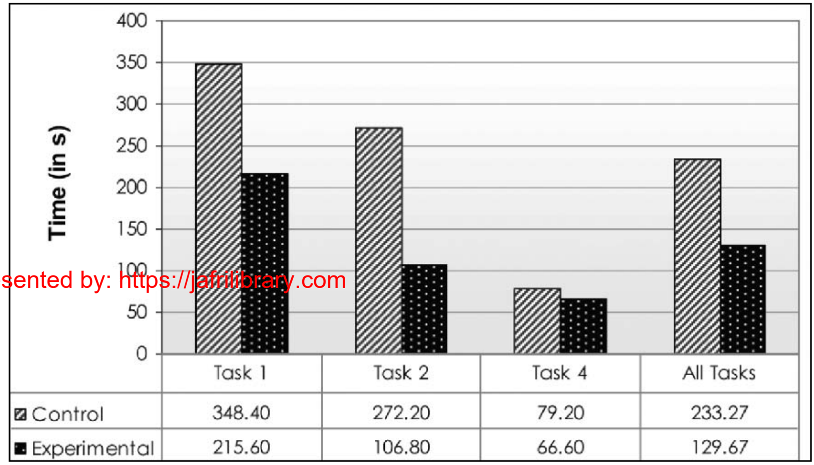

Figure 6: Average Task Times for Each Group in Main Study

From the above graph it is clear that the Experimental group performed the tasks significantly faster than the Control group with the Experimental group completing tasks on average 44.4% faster than the Control group. As already explained, these figures do not include Task 3 which was the logging task. The reason for this is that the task was of a fixed length, i.e. 302 seconds. As such, including this task would achieve nothing other than inflating the group times of each group by the same amount.

An interesting issue is raised by the results for Task 4. When each task time was analysed individually, two-sided Wilcoxon-Mann-Whitney exact tests yielded p-values of 0.0079 for both Task 1 and Task 2. However, the p-value for Task 4 was 0.254, indicating that there was no statistically significant difference between the average times for each. Indeed, Figure 6 clearly shows that while there is a difference between the two groups, the difference is not as pronounced as in the other tasks. There are several possible explanations for this. Firstly, it is possible that analysing each task time individually resulted in a Type I error or spurious result.

Another, more probable, explanation is that, bearing in mind this task involved selecting a method for saving a completed log file to a specific location, the task is not as complex as the other tasks and as such there is a smaller likelihood of serious problems arising. We could argue that this task had more to do with users' knowledge of the Windows operating system than their knowledge of DigiLog and as such, this – and not the user guide – played a greater role in the outcome. Nevertheless, analysing the

overall time taken by each group to complete all tasks, we can see that the Experimental group performed the tasks faster than the Control group.

An interesting issue raised by the results for the main study is that the times are significantly lower than those observed in the pilot study. This can be attributed to the fact that, in comparison to the pilot group, the participants in the main study were more experienced in the particular work environment, had a greater understanding of the context in which the software would be used and used computers every day. While the pilot group were a closer match to the profile of real users, the participants in the main study were an even better match.

Error Rates

A two-sided Wilcoxon-Mann-Whitney exact test of the null hypothesis of no difference between the total error score for the Control group and Experimental group was also conducted. This yielded a *p-value of 0.0079* indicating that there was a statistically significant difference between the scores for the Control group and the Experimental group.

We can see from Figure 7 that the Control group using the original version of the user guide committed on average 31.9 errors in comparison to the Experimental group who committed 10.9 errors. This represents quite a significant divergence in the effectiveness of each group of users in performing tasks. Indeed, looking at individual error criteria we can see some rather striking trends.

- Some 60% of Control group failed to complete one task (criterion 1)
- During the tasks, the Experimental group consulted the user guide 44.4%
- less than the Control group (criterion 2)
- The Experimental group remembered 84.2% more commands and icons than the Control group (criterion 3)

- The Experimental group made 69.2% fewer incorrect icon choices than the Control group (criterion 4)
- The Experimental group made 73.5% fewer incorrect menu choices than the Control group (criterion 5)
- The Experimental group did not ask any questions during the test whereas the Control group asked on average 1.8 questions (criterion 6)
- There were no observations of frustration in the Experimental group while in the Control group there were three individual observations of frustration (criterion 7)

- There were 76.9% fewer observations of confusion in the Experimental group (criterion 8)
- There was a single observation of satisfaction in the Experimental group compared with no observations in the Control group (criterion 9)
- The Control group used 58.8% more incorrect commands or input (criterion 10)
- In the Control group, two participants mistakenly thought they had completed one or more tasks; in the Experimental group, all participants successfully completed the tasks before finishing work (criterion 11)

Comparing the average total error rates for each group (Figure 8), we can see that the Experimental group, which used the user guide featuring Iconic Linkage, committed on average 10.9 errors over the course of the test. In comparison, the Control group, which was using a user guide without Iconic Linkage, committed on average 31.9 errors over the course of the test. To put this another way, we can say that the Experimental group made 65.8% fewer mistakes than the Control group.

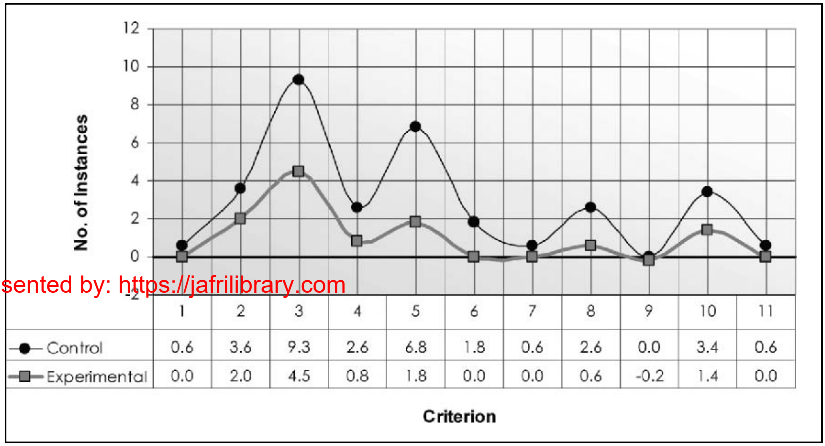

Figure 7: Average Usability Criteria Error Scores for Each Group in Main Study

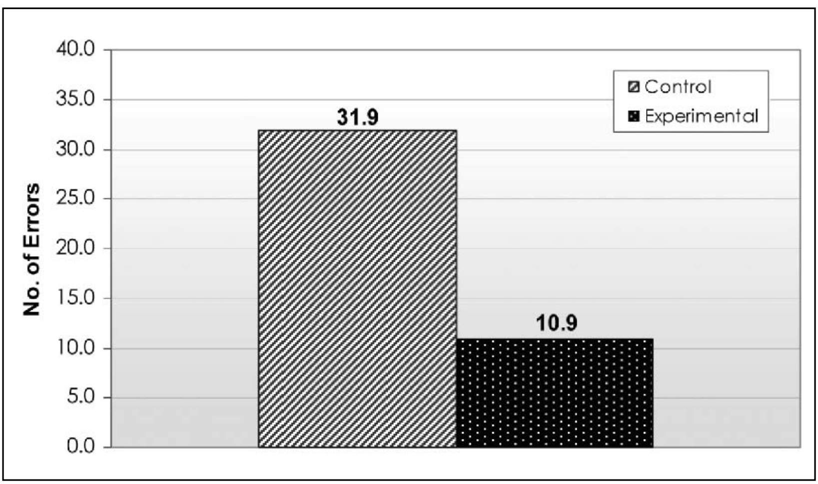

Figure 8: Average Group Error Rates in Main Study

Results of QUIS Usability Survey

A two-sided Wilcoxon-Mann-Whitney exact test of the null hypothesis of no difference between the average overall score for the Control group and Experimental group was conducted. This yielded a *p-value of 0.0079*

indicating that there is a statistically significant difference between the average scores across the two groups. On the basis of the data collected, members of the Experimental group were, on average, 24.3% more satisfied with the user guide than the Control group. Examining the graph of average responses from the QUIS questionnaire (Figure 9), however, it immediately becomes apparent that there are a number of areas where there is no major difference visible between the two groups. The following paragraphs will examine these areas.

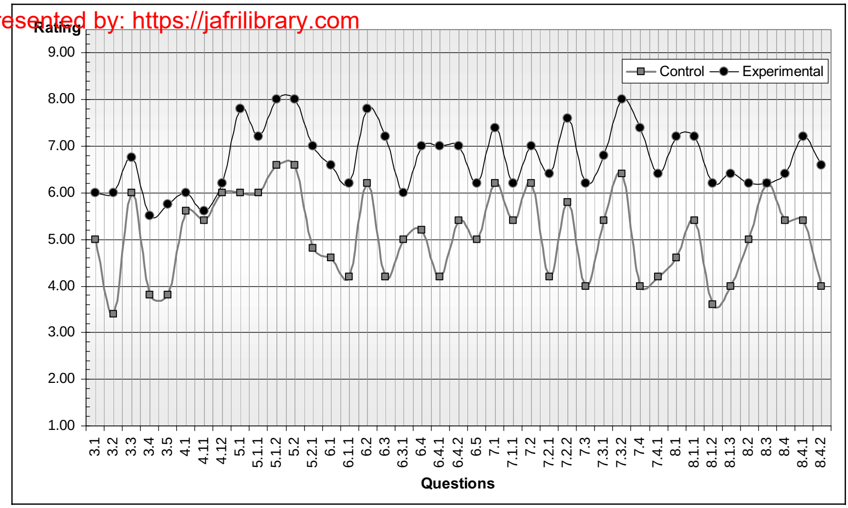

Figure 9: Average Responses in User Satisfaction Survey

Questions 4.1, 4.11 and 4.12

- Page layout was helpful: never/always
- Amount of information displayed on a page: inadequate/adequate
- Arrangement of information on page: illogical/logical

The results for these questions converge between the two groups. The fact that there is no noticeable difference between the two groups can be explained in light of the fact that the questions relate to the amount of information on the page and the layout of information on the page. Neither of these factors was altered in the study. In fact, every effort was made to ensure that these factors were identical in both versions.

Question 8.3

- Remembering names and use of commands: difficult/easy

This question relates to how easy subjects found it to remember the names of commands and how to use them. Both groups provide almost identical ratings indicating that they both felt that they could remember this information with above average ease. However, it is clear from the Post-Test Survey (PTS) scores for the two groups that this is most certainly not the case. With an average PTS score of 9.3, the Control group clearly found it more difficult to remember commands than the Experimental group with a PTS score of 4.5. To put this in context, Control subjects remembered on average just 5.7 commands whereas subjects in the Experimental group remembered on average 11.5 commands; the Experimental group performed roughly twice as well as the Control group.

In the other sections of the questionnaire, we can see quite striking differences in the satisfaction levels of both groups with regard to the various aspects of the user guide, product and tasks. In the following paragraphs we will examine these sections.

Section 3: Overall User Reactions to System

terrible/wonderful: The Control group gave an average rating of 5. The Experimental group were 20% happier with average rating of 6.

frustrating/satisfying: On average, the Control group found it frustrating, awarding it an average rating of 3.4. The Experimental group were approximately 76.4% more satisfied with an average rating of 6.

dull/stimulating: The Control group found the system stimulating overall, awarding it an average rating of 6. Experimental group were 12.5% happier awarding it an average rating of 6.75.

difficult/easy: With an average rating of 5.5, the Experimental group regarded the system almost 44.7% easier to use than the Control group which gave an average rating of 3.8.

rigid/flexible: The Control group felt that the system was more rigid than the Experimental group with average ratings of 3.8 for the Control group and 5.75 for the Experimental group.

While these criteria do not directly relate to the incidence of Iconic Linkage (IL) in the user guide, they do show that the increased usability caused by the inclusion of IL results in improved user attitudes towards a product. In this case it is clear that users have a better opinion of the software simply because the user guide is more usable.

Section 5: Terminology

This section deals with the terminology used in the user guides. On the whole, both groups felt that the level of terminology used was appropriate and consistent. The Control group did, however, feel that computer terminology was used slightly too frequently. Bearing in mind that both user guides contained exactly the same information and that if anything, terms were more likely to be repeated in the Experimental version, it is difficult to understand how this perception arises. We might presume that the overall difficulties encountered by subjects resulted in "spill-over" into other areas of the user guide with the result that difficulties in one area can result in lower levels of satisfaction throughout the user guide in a form of "guilt by association" effect.

Section 6: User Guide

For *Question 6.1* "The user guide is: confusing/clear" the average Control group rating was 4.6 while the average Experimental group rating was 6.6. This indicates that the Experimental group found the user guide 43.5% clearer than the Control group did.

In response to *Question 6.1.1* "The terminology used in the user guide: confusing/clear", the average Control group response was 4.2 and the average Experimental group response was 6.2. Despite the terminology in both versions being almost identical, these figures represent a 47.6% improvement in the Experimental group. One possible explanation for this is the effect on users' overall attitudes to the document caused by improved usability and accessibility of information.

Question 6.2 "Language used in the user guide is consistent: never/always" provided some interesting results in terms of how noticeable the repetition of textual structures was perceived by subjects. The average Control group response to this question was 6.2 and the average Experimental group response was 7.8. While on the one hand this shows that the Experimental version of the user guide was more consistent than the Control version, the fact that there was no large increase in consistency from the point of view of the user indicates that users were not particularly aware of the repetition of textual structures caused by IL. As such, we can argue that IL is not as noticeable and, therefore, not as disruptive to users as one would imagine. On a related issue, however, one subject in the Control group mentioned in the post-task interview during the pilot study, without any form of prompting or specific questioning, that the lack of consistency and Iconic Linkage (the subject did not use this term but rather described the phenomenon) in the user guide proved problematic and distracting and hindered comprehension. Although no other subject mentioned this (the

information was in no way solicited) it does highlight the fact that while the presence of Iconic Linkage is not always detected by readers, a lack of Iconic Linkage may be noticed.

In response to *Question 6.3* "Understanding information in the user guide is: difficult/easy" the Experimental group, with an average rating of 7.2, found the information in the user guide 71.4% easier to understand in comparison to the Control group which responded with a rating of 4.2.

For *Question 6.3.1* "Finding a solution to a problem using the user guide: impossible/easy" the Experimental group with an average rating of 6, found it on average 20% easier to find solutions to problems than the Control group with a rating of 5.

In responding to *Question 6.4* "Amount of help given: inadequate/adequate" the Control group gave an average rating of 5.2 and the Experimental group gave a rating of 7. From these figures we can see that the Experimental group were 34.6% more satisfied with the information provided by the user guide even though both versions contained the exact same information. Again, we can see that improved usability in one area can improve overall subjective ratings in other areas.

In *Question 6.4.1* "User guide defines specific aspects of the system: inadequately/adequately" with an average rating of 7, the Experimental group found the user guide's definition of specific aspects of the system 66.6% better than the Control group who gave an average rating of 4.2.

In response to *Question 6.4.2* "Finding specific information using the user guide: difficult/easy", the Experimental group with a rating of 7 found it 29.6% easier to find specific information using the user guide than the Control group which gave an average rating of 5.4.

On the basis of the responses from the two groups to *Question 6.5* "Instructions for performing tasks are clear and unambiguous: never/always", the Experimental group (6.2) found the instructions 24% clearer and easier to understand than the Control group (5).

Section 7: User Guide Content & Structure

In response to *Question 7.2.1* "Information for specific aspects of the system was complete and informative", the Control group gave an average rating of 4.2 compared to the Experimental group which gave a rating of 6.4. While both user guides had identical information, the apparent difficulty encountered by subjects in finding information would suggest that they had problems assimilating or recognising information and as such they could not recognise the information content of what they were reading.

For *Question 7.2.2* "Information was concise and to the point" the average Control group rating was 5.8 compared to the Control group's rating of 7.6. The average ratings for the Control group can be interpreted as meaning that the presentation of information in the Control user guide (despite the slightly higher word count) was not substandard, i.e. it was of an acceptable standard. The Experimental user guide, however, was more than acceptable in comparison.

For *Question 7.3* "Tasks can be completed: with difficulty/easily" the average rating for the Control group was 4 while the Experimental groups average rating was 6.2. This represented a 55% improvement in the Experimental group's rating.

This result is borne out by task completion figures for both groups. In the Control group, three out of the five subjects (60%) failed to complete all of the tasks. In comparison, all of the subjects in the Experimental group completed all of the tasks.

For *Question 7.3.1* which asks whether subjects found the instructions for completing tasks clear or confusing, the average response for Experimental group was 6.8 in contrast to 5.4 for the Control group. The Experimental group, therefore, found the instructions 25.9% clearer.

For *Question 7.4*, "Learning to operate the system using the user guide was: difficult/easy" the Experimental group gave an average rating of 7.4 compared to the average Control group rating of just 4. This represents quite a drastic difference with the Experimental group finding it 85% easier to learn to use the system.

Question 7.4.1 "Completing system tasks after using only the user guide was: difficult/easy" provided similar results with the Control group giving an average rating of 4.2 compared to 6.4 given by the Experimental group. Again, these figures are echoed in the error rates for each group and the task completion rates.

Conclusions

In this chapter we have established that in order to test the hypothesis that Iconic Linkage improves the usability of software user guides, some form of summative evaluation involving users is essential. With our definition of usability consisting of both quantifiable and subjective components, the need to collect both quantitative and qualitative data is apparent.

This chapter discussed the various methods for collecting data and the ways in which they can be implemented. It is clear from this chapter that indirect observation is preferable over direct observation methods because of the risk of influencing subjects. However, the nature of the tasks and the facilities available for setting up a usability laboratory made it impossible to conduct the experiment without the administrator being present in the laboratory. While this is less than ideal, the effect of the administrator's presence was minimised through careful positioning and regulated interactions during the experiments.

We also considered a variety of data collection methods such as interviews, video and audio recording, screen logging and questionnaires. After discussing each of these in detail, it was possible to select and reject methods on the basis of their suitability for the requirements of this study.

An examination of literature relating to previous experiments and case studies was carried out in the hope of finding useful information on conducting usability experiments. What emerged from this review is that there seems to have been a shift away from documentation usability testing over the past decade or so, particularly with regard to print documentation. Those that do deal with documentation, regard documentation as including both print and online texts. Other studies which exhibit certain compatibilities with this study often differ in terms of their goals and objectives or they make inappropriate assumptions. Nevertheless, by analysing several studies, it was possible to extract useful pointers for conducting a usability experiment. Of the literature reviewed, only two stand out as being particularly relevant or useful. These studies were discussed in detail.

With this background knowledge, the chapter proceeded to describe the preparations, procedures and results of a pilot study conducted to test the methodology and protocols for the study. This consisted of producing materials and forms, recruiting participants, defining evaluation criteria for testing the user guide etc. The chapter describes the problems encountered during the pilot study. One such problem which emerged related to the specification of evaluation criteria. This proved problematic because although certain criteria may be useful or important, they may not necessarily be measurable due to the nature of the product and tasks. Similarly, data collection tools and methods are not always suitable for recording a particular type of information. Consequently, a number of changes had to be made before conducting the main study.

Other issues such as those encountered with the participant who refused to follow instructions show that a great deal of preparation, flexibility and discipline on the part of the tester are essential in order to cope with

unforeseen eventualities. Future research of this nature would need to take into account the notion of field-dependent and field-independent people, i.e. people who are more or less likely to use a user guide to learn how to use software. Screening of participants would, therefore, need to identify whether potential participants tend to use user guides or whether they prefer to "figure it out for themselves". This could be done simply by asking them how they normally learn how to use software or by using the *Group Embedded Figures Test* – GEFT (Witkin *et al.* 1971).

Having described the pilot study and the modifications made to the test procedures, the chapter proceeded to discuss the main study. From this we can see quite clearly that Iconic Linkage clearly has a positive effect on the usability of software user guides.

Taking the first of the three components of usability, i.e. the speed with which users perform tasks, the results of the empirical study clearly show that subjects using a user guide into which Iconic Linkage has been introduced performed tasks significantly faster than those using a user guide with no Iconic Linkage.

An interesting issue arises in relation to the results of the pilot study and the main study with regard to task times. While both studies showed the same dichotomy of results between the two groups, the subjects in both groups in the main study performed the tasks significantly more quickly than their counterparts in the pilot study. This can be explained by the fact that the participants in the main study were more experienced in the type of work involved, had more experience of using computers and had a better understanding of the context in which the software is used than those in the pilot study.

With regard to error rates for the two groups in the main study, the results show that the Control group made three times more mistakes than the Experimental group. Considering this more closely we can see that the Experimental group completed more tasks, worked more efficiently and made fewer mistakes using icons and commands. The Post Test Survey also shows that the Experimental group remembered more information about the software than the Control group.

In assessing the final component of our definition of usability, the user satisfaction questionnaire shows that attitudes to and satisfaction levels with the software were considerably more favourable in the Experimental group than in the Control group. The Experimental group found that the user guide which featured Iconic Linkage was clearer, easier to understand and more effective in helping them achieve their goals than the Control group. Interestingly, questions in the questionnaire designed to detect whether users

detected the presence of Iconic Linkage indicated that the Experimental group did not detect Iconic Linkage. Both groups gave broadly similar ratings for the consistency and amount of repetition. This indicates that introducing repetition into a user guide does not necessarily represent a distraction for users. Indeed, one subject in the Control group actually commented that the lack of consistency in phrasing instructions was problematic for comprehension. No other subject mentioned this but this is a definite reference to Iconic Linkage – the user said that the lack of Iconic Linkage was distracting and resulted in the need to "refocus" after each sentence.

In the main study, another interesting issue arose in relation to interference between users' existing knowledge and the new information they were trying to learn. The existing knowledge domain in question related to users' prior knowledge of *Microsoft Word* and *QuicKey*. One member of the Control group noted that the way *QuicKeys* worked in a fundamentally different way to the way *Microsoft Word* implements a similar function. Thus, prior knowledge hampered users' learning of the new information. This information was not specifically requested, it was volunteered by the subject; no other subjects mentioned this.

Overall, the empirical study shows that Iconic Linkage is a viable and effective strategy for improving the usability of translated software user guides. Introducing Iconic Linkage into a text during translation makes it easier for users to understand the information and helps them learn to use the software more quickly. Even with the small sample sizes used in this study, it has been shown with a high level of statistical reliability that Iconic Linkage makes user guides more effective, that users can perform their tasks more quickly and retain more knowledge for longer.

The fact that clear improvements were detected across the three components of the test, i.e. task times, error rates and satisfaction levels, also shows that concerns regarding the possible confounding influence of the slight improvement in readability or the elimination of passive sentences are unfounded. Even if, as was discussed on page 215, the absence of passive sentences was the only improvement in the user guide, the improvements would be restricted to the task times alone because participants would have been able to read the user guide more quickly. However, because error rates and satisfaction levels also improved among the Experimental group, it is unlikely that this is simply due to the lack of passives. As such, it is difficult to treat the elimination of passives as a genuine confounding factor.

Chapter 6
Conclusions

The preceding chapters have explored a wide range of theoretical and practical knowledge in areas such as translation, technical communication, psychology and usability engineering in the hope of shedding light on ways in which technical translation can be studied, practiced and improved. From this series of discussions and examinations of various areas, it is clear that technical translation is much more complex and interesting an activity than many would initially think.

The assertion in Chapter 1, for example, that the success or failure of technical translations depends on much more than just specialised terminology and that a text's style, register, content, communicative function, usability and so on all play a role makes technical translation both interesting and demanding. Our role as communicators means that we have a responsibility to our readers to ensure that the information is communicated effectively. This is particularly true in technical and instructional texts. The variety of legislation relating to technical documentation and translation makes our task more challenging while at the same time providing clear evidence of the need to improve our "products" and they way we produce them. Looking at technical translation as a communicative activity rather than as a purely text-based, linguistic activity opens the door to a variety of new and interesting avenues of investigation.

From a theoretical point of view, it could, of course, be argued that the approaches and attitudes expressed in this book are steps in the wrong direction, i.e. away from translation in its purest form. This is particularly true in the context of certain translation theories, approaches and genres such as Reiss, House or the various linguistic approaches to translation. Such a criticism may or may not hold true depending on your own personal views of translation and translation theory. But looking at translation through the lens of Skopos theory and, to a lesser extent, relevance theory, it is difficult to argue against engineering translations to ensure maximum acceptability, effectiveness, usability and efficiency from the target audience's point of view. In the context of Skopos theory, we see the need to take the needs of the initiator and reader into account when producing translations. Relevance

theory with its emphasis on the cognitive aspects of texts, introduces us to the idea of shaping translations to suit certain cognitive and psychological aspects of readers. Admittedly, neither approach to translation is without its problems and deficiencies but they do help us to broaden our field of vision over the translation landscape. They help show that translation does not necessarily need to kowtow to the supremacy of the source text and be subject solely to the myriad relationships between source and target text. Instead, by adopting a more communicative, professionally-orientated stance to translation we should realise that translation is a service (hopefully) provided by professionals for customers who place their trust in the skills, abilities and motivations of the translators they engage to provide a service for them.

In adopting an approach which is unashamedly aimed at the target language and target audience, this book has shown that by examining the circumstances under which technical texts are produced in the target language, we can gain a better understanding of why texts are produced and what they should do when translated into the target language. Of course, this clear target-orientation will undoubtedly set all manner of alarm bells ringing for some people critical of the fact that there appears to be scant regard for some sacred, yet excruciatingly nebulous notion of "faithfulness" to the source text and author. It would be inviting trouble to say that there should not be **some** link between the source and target texts but in certain cases, particularly in technical translation, it would be foolish and indeed grossly negligent for a translator to put anything other than the needs of the target audience first with all of the implications this has for the role of the source text. This is not because we are bound for reasons of good business to appease our paymasters, but because under normal circumstances our ethical, moral and legal obligations rest with the target audience; they are, after all the intended receivers in the communicative act. As facilitators of communication, this is where our attention must be focussed.

Cognetics and Iconic Linkage

At first glance it may be difficult to see how a strategy such as Iconic Linkage can improve translations. After all, it involves is making a text more repetitive, and possibly even more boring, by removing variation and creativity from a source text and replacing it with standardised, invariant and simplified formulations. So how can it be regarded as a valid translation strategy let alone as a way of actually improving the quality of translations? Indeed, such a notion would appear to run contrary to conventional logic

and accepted wisdom in translation studies. The justification for this is quite simple if we think back to the primary aim of technical texts. In dealing with texts of a technical nature we are concerned first and foremost with the payload of a text, i.e. its content, and how well it can be assimilated by the reader. Consequently, any additional sub-functions or features of the text are of lesser or negligible importance. With this in mind, the idea of making a text less entertaining or fun to read is of little consequence when we consider that the text was never intended to be entertaining or fun; it was only ever intended to convey information effectively. In any case, none of the subjects in the usability study regarded the edited user guide as monotonous or boring beyond what is normal for a text of this type. With our attention firmly fixed on the target audience and their needs, it is clear that the need to extract information from a text is of greater importance than to enjoy reading it. But aesthetic considerations aside, the clear benefits of this strategy for the reader in terms of clarity, usability and learnability surely outweigh any perceived stylistic shortcomings.

Future Directions

Despite its prominence in this book, Iconic Linkage is just one way of adapting translations to suit the cognitive abilities of readers. It may be one of the more drastic measures but it is still just one possible strategy. There are numerous conceivable ways in which texts can be engineered to improve usability, some of which are quite simple, for example, making sure sentences are not excessively long or that we phrase instructions logically. As mentioned in Chapter 4, Iconic Linkage is not suitable for use in all texts as it relies on a certain level of repetition of concepts and information for it to be practical and effective.

The challenge now is to examine other ways of improving the usability of texts not just in technical genres but in other areas. It should be pointed out here that while this study was concerned with software user guides, there is no reason why Iconic Linkage could not be used to produce the same benefits in any kind of instructional text dealing with any subject. It would also be interesting to see whether translations of poems or plays could incorporate some aspects of cognitive psychology to trigger specific responses or thought processes or to inspire specific mental images in readers. Another possible application of cognitive engineering would be to see whether persuasive texts such as advertisements could be made *more persuasive* thanks to a translator's ability to engineer the translation. Could we make legal texts more usable through translation despite the very strict

conventions and requirements that affect it? Could usability play a role in any other types of translation?

Another area of potential interest both academically and professionally would be to examine the role of controlled languages on the usability of texts. One of the aims of Iconic Linkage is to restrict the amount of variation which can occur in the way identical information is expressed in a text. While this is admittedly rather simplistic in comparison, it is hard to ignore the similarities with the aims of controlled language. By specifying specified ways in which information can be expressed in a text, controlled language seeks to make texts easier to understand and/or easier to translate using translation memory (TM) or machine translation (MT) systems depending on the particular controlled language. It is conceivable that using a comprehensive controlled language we can implement the basic notion of Iconic Linkage in a wider range of texts and achieve similar improvements in usability while at the same time improving translatability.

The concept of Iconic Linkage also highlights new areas for research into TM tools. We saw how one such tool, *Trados Translators' Workbench*, was used to introduce Iconic Linkage in one version of the user guide. However, this process was not without its difficulties, chiefly because of the tool's inability to identify semantically identical sentences. While TM tools can identify Iconic Linkage in a text, and indeed benefit immensely from it (is designed to work best when there are identical sentences in a text), an ideal situation would see TM tools play more of a role in introducing Iconic Linkage into a text in the first place. To do this, the technology used in TM tools would need to be re-evaluated to see how they can "understand" text in order to identify semantically identical information.

Iconic Linkage also has an impact on the business of translating user guides. Since Iconic Linkage essentially involves introducing consistency and repetition into a text and limiting the ways in which information can be phrased, it has a significant effect on the effectiveness of TM tools. These tools are essentially databases which are used to store sentences and their translations so that if the same sentence is encountered again during transla-

tion, the existing translation is retrieved from the database and inserted into the text. Since Iconic Linkage increases the likelihood that a text will contain an increased proportion of repetitions, the effectiveness of TM tools will be increased. The benefits of this to industry include faster turnaround times for translations, greater reusability and consistency of translations. And because repetitions reduce the total number of words to be translated, the overall translation costs are reduced. Add to this the improved usability introduced by Iconic Linkage and we can see how beneficial Iconic Linkage could be to the translation industry.

It is also clear from the preceding chapters just how important it is for translators to be familiar with technical communication in general. As mentioned earlier, if we want to translate technical texts effectively and achieve the highest standards of quality it is essential that we understand how these texts are produced in the first place. Technical translations invariably compete with original texts in the target language. As such, translations must be produced to the same standards and expectation as "native" texts. This leads us to question whether technical communication should be included as part of a translation programme. Already several institutions include modules on technical writing as part of their translator training programmes but it is clear that they are the exception rather than the rule.

In short, technical translation can benefit immensely from technical communication in its broadest sense: it helps us produce texts that can compete with original target language texts; it helps us engineer texts to meet the quality goals of acceptability, accuracy and usability and it provides us with mechanisms for assessing their quality. It is essential that we regard technical translation as process where information is designed and engineered rather than as a purely textual process.

Bibliography

Alcaraz, E. and B. Hughes (2002) *Legal Translation Explained*. Manchester: St. Jerome Publishing

Amman, M. and H.J. Vermeer (1990) *Entwurf eines Curriculums für einen Studiumgang Translatologie und Translatorik*. Heidelberg: Institut für Übersetzen und Dolmetschen

Anderson, J.R. (2000) *Cognitive Psychology and its Implications*. 5th edition. New York, USA: Worth Publishers

Apple Computer Inc. (1978) *Apple II BASIC Programming Manual, Informative and Entertaining Reading for Novice and Professional Alike*. Press release dated 8th June, 1978 [online] Available from: http://www-sul.stanford.edu/mac/primary/docs/pr4.html [Accessed: 16/04/02]

Asch, S.E. (1956) Studies of Independence and Conformity: A Minority of One against a Unanimous Majority. In: *Psychological Monographs*, 70, pp. 1-70

Austin, M. and R.F. Dodd (1985) *The ISTC Handbook of Technical Writing and Publication Techniques: A Practical Guide for Managers, Engineers, Scientists and Technical Publications Staff*. London, UK: William Heinemann Ltd.

Bannon, L. (1991) From Human Factors to Human Actors The Role of Psychology and Human-Computer Interaction Studies in Systems Design. In: J. Greenbaum and M. Kyng (eds) *Design at work: Cooperative Design of Computer Systems*. Hillsdale, USA: Lawrence Erlbaum Associates, pp. 25-44.

Bødker, S. (1991) *Through the interface: A human activity approach to user interface design*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Borowick, J.N. (1996) *Technical Communication and its Applications*. New Jersey, USA: Prentice-Hall

Broadbent, D. (1958) *Perception and Communication*. London, UK: Pergammon Press

Brockman, R.J. (1990) *Writing Better Computer User Documentation: From Paper to Hypertext – Version 2.0*. New York, USA: John Wiley & Sons

Buchanan, R.A. (1992) Textbase Technology: Writing with Reusable Text. In: P. O' Brian Holt and N. Williams (eds) *Computers and Writing: State of the Art*. Oxford: Intellect Books

Budin, G. (2000) *Terminology – Translation – Technical Communication*. [online] Available from: http://mailbox.univie.ac.at/~buding8/dint.htm [Accessed: 21/03/04]

Byrne, J. (1999) International English: Making A Case for Efficient and Clear Communication. In: *ATA Chronicle*, Vol. XXVII, No. 10

Card, S.K., T.P. Moran A. and Newell (1983) *The Psychology of Human-Computer Interaction*. New Jersey, USA: Lawrence Erlbaum Associates

Catford, J.C. (1965) *A Linguistic Theory of Translation*. London: Oxford University Press

Chambers (1993) *Chambers Maxi Paperback Dictionary*. Edinburgh: W&R Chambers

Cherry, E.C. (1953). Some experiments on the recognition of speech, with one and with two ears. In: *Journal of the Acoustical Society of America*, 25, pp. 975-979

Coe, M. (1996) *Human Factors for Technical Communicators*. New York, USA: John Wiley & Sons

Cognitas (2003a) *Gute Bedienungsanleitung – geringeres Haftungsrisiko*. cognitas - Gesellschaft für Technik-Dokumentation: Paderborn [online] Available from: http://www.tuev-sued.de/presse/pressearchiv/iiftqseozlfy.asp [Accessed: 21/03/04]

Cognitas (2003b) *DOCcert-Kooperationsvertrag bietet deutschen Herstellern erhöhten Schutz vor Bedienungsanleitungs-Reklamationen*. cognitas - Gesellschaft für Technik-Dokumentation: Paderborn [online] Available from: http://www.tuev-sued.de/start/pressearchiv/jtusowqgxxza.asp. [Accessed: 21/03/04]

ComputerBild (1999) Holen Sie sich beim Händler Geld zurück. In: *ComputerBild* 25/99, pp. 12-17

Council of the European Union (1993) *Council Directive 93/42/EEC of 14 June 1993 concerning medical devices*. Official Journal of the European Communities No. L N169/1 of 1993-07-12)

Council of the European Union (1998a) *Council Resolution of 17 December 1998 on operating instructions for technical consumer goods*. Official Journal of the European Communities (98/C 411/01)

Council of the European Union (1998b) *Directive 98/37/EC of the European Parliament and of the Council of 22 June 1998 on the approximation of the laws of the Member States relating to machinery*. Official Journal of the European Communities L 207/1

Curwen, C. (2002) *Comments on "Give them Printed Documentation, too!!!"*. [online] Available from: http://www.tc-forum.org/topicus/ru31comm.htm. [Accessed: 22/05/2004]

D'Agenais, J. and J. Carruthers (1985) *Creating Effective Manuals*. Cincinnati, USA: South Western Publishing Co

Deutsch, M. and H.B. Gerard (1955) A Study of Normative and Informative Social Influences on Individual Judgements. In: *Journal of Abnormal and Social Psychology*, 51, pp. 629-636

Dillon, A. (1992) Reading from paper versus screens: A critical review of the empirical literature. In: *Ergonomics*, 35(10), pp. 1297-1326

Dix, A. (1998) *Human-computer Interaction*. 2nd Edition. New Jersey, USA: Prentice Hall

Downton, A. (1991) *Engineering the human-computer interface*, London; New York, USA: McGraw-Hill Book Co

Dumas, J.S. and J.C. Redish (1999) *A Practical Guide to Usability Testing*. Exeter, UK: Intellect Books

Ehn, P. and J. Löwgren (1997) Design for Quality-in-use: Human-Computer Interaction Meets Information Systems Development. In: M. Helander, T.K. Landauer, P. Prabhu (eds) *Handbook of Human-Computer Interaction* (2nd edition). The Netherlands: Elsevier Science. pp. 299-313

Ellis, H.C. and R.R. Hunt (1993) *Fundamentals of Cognitive Psychology*. 5th edition. Dubuque, Iowa: WCB Brown & Benchmark Publishers

EN 292-2:1991 - Safety of machinery. Basic concepts, general principles for design. ISBN: 0580203646

EN 62079:2001 - Preparation of instructions. Structuring, content and presentation. ISBN: 0 580 38240 0

Esselink, B. (2000) A Practical Guide to Localization. The Netherlands and Philadelphia, USA: John Benjamins

Eysenck, M.W. (2000) *Psychology: A Student's Handbook*. Sussex & New York: Psychology Press

Fabrizio, R., L. Kaplan and G. Teal (1967) Readability as a function of the straightness of right-handed margins. In: *Journal of Typographic Research*, 1, pp. 90-95

Faulkner, C. (1998) *The Essence of Human-Computer Interaction*. London, UK and New York, USA: Prentice Hall

Fawcett, P. (1997) *Translation and Language: Linguistic Theories Explained*. Manchester: St. Jerome Publishing

Fishbach, H. (1998) Guest Editor's Preface. In: *Translation and Medicine*, American Translators Association Series, Vol. X, pp. 1-12

Foss, J.D., M.B. Rosson and P.L. Smith (1981) Reducing Manual Labor: An Experimental Analysis of Learning Aids for a Text Editor. In: *Proc. Human Factors in Computer Systems*. Washington: ACM. pp. 332-336

Franco Aixelá, J. (2004) The Study of Technical and Scientific Translation: An Examination of its Historical Development. In: *Journal of Specialised Translation*, Issue 1, pp. 29-49

Fredholm, L. (2001) Article on Ivan Petrovich Pavlov. [online] Available from: http://www.nobel.se/medicine/educational/pavlov/readmore.html [Accessed: 24/10/02]

Gavin, H. (1998) *The Essence of Cognitive Psychology*. Hertfordshire, UK: Prentice Hall Europe

Gerzymisch-Arbogast, H. (1993) Contrastive Scientific and Technical Register as a Translation Problem in Scientific & Technical Translation. In: S.E. Wright (ed) *Scientific and Technical Translation*: American Translators' Association Scholarly Monograph Series, Vol. VI 1993. Amsterdam, The Netherlands and Philadelphia, USA: John Benjamins

Ginsburg, H.P. and S. Opper (1988) *Piaget's Theory of Intellectual Development*. 3rd edition. New Jersey, USA: Prentice Hall

Goldsmith, D. (1995) *Einstein's Greatest Blunder?* Cambridge, MASS.: Harvard University Press

Gommlich, K. (1993) Text typology and translation-oriented text analysis. In: S.E. Wright (ed) *Scientific and Technical Translation*: American Translators' Association Scholarly Monograph Series, Vol. VI 1993. Amsterdam, The Netherlands and Philadelphia, USA: John Benjamins

Göpferich, S. (1993) Die translatorische Behandlung von Textsortenkonventionen in technischen Texten. In: *Lebende Sprachen*. No. 2/93, pp. 49-52

Gregory, M. and E.E. Poulton (1970) Even versus uneven right-margins and the rate of comprehension in reading. In: *Ergonomics*, 13, pp. 427-434

Gutt, E.A. (1991) *Translation and Relevance: Cognition and Context*. Oxford, UK: Basil Blackwell

Halverson, S. (1997) The Concept of Equivalence in Translation Studies: Much Ado About Something. In: *Target*, 9(2), pp. 207-233

Harrison, S. and P. Mancey (1998) Optimal amount of time for obtaining accurate usability-test results. In: *ACM Proceedings of the 16th annual international conference on Computer documentation*. New York, USA: ACM Press. pp. 173-179

Hartley, J., and P. Burnhill (1971) Experiments with unjustified text. In: *Visible Language*, 5, pp. 265-278

Hatim, B. (2001) Teaching and Researching Translation. London: Longman

Heino, H. (1992) Spezialprobleme beim Übersetzen fachsprachliche Texte. In: *Fachsprache*, Vol. 3-4, pp. 111-120

Helsel, D.R. and R.M. Hirsch. (1992) *Statistical Methods in Water Resources*. Amsterdam, The Netherlands: Elsevier Publishers

Hill, S. (1995) *A practical introduction to the human-computer interface*. London, UK: DP Publications

Holz-Mäntärri, J. (1984) *Translatorisches Handeln: Theorie und Methode*. Helsinki: Suomalainen Tiedeakatemia

Hönig, H. (1997) Positions, Power and Practice: Functionalist Approaches and Translation Quality Assessment. In: *Translation and Quality*. Christina Schäffner (ed.), Clevedon: Multilingual Matters, 1998, pp. 6-34

Houghton-Alico, D. (1985) *Creating Computer Software User Guides: From Manuals to Menus*. New York, USA and London, UK: McGraw-Hill

House, J. (1977) *A Model for Translation Quality Assessment.* (1st edition) Tubingen: Gunter Narr Verlag

House, J. (1981) *A Model for Translation Quality Assessment.* (2nd edition) Tubingen: Gunter Narr Verlag

Hukins, D. (1981) *X-Ray Diffraction by Disordered and Ordered Systems*. Oxford: Pergamon Press

ISO 9241-11:1998 Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability. Geneva: International Organization for Standardization

Jung, M. and O. Becker (2003) *Die Bedienungsanleitung - kundenfreundlich und rechtssicher*. cognitas - Gesellschaft für Technik-Dokumentation: Paderborn [online] Available from: http://www.cognitas.de/pdf/tuev_cognitas.pdf. [Accessed: 21/03/04]

Kade, O. (ed.) (1977) *Vermittelte Kommunikation, Sprachmittlung, Translation*. Leipzig: VEB Verlag Enzyklopädie

Kalat, J.W. (2004) *Biological Psychology*. 8th Edition/International Student Edition. London: Thomson Wadsworth

Kellogg, R.T. (1988) Attentional overload and writing performance: effects of rough draft and outline strategies In: *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Vol. 14, No. 2, pp. 355-6

Kingscott, G. (2002) Technical Translation and Related Disciplines. In: *Perspectives: Studies in Translatology*. Vol. 10:4, pp. 247-255

Kirakowski, J. (2000) *Questionnaires in Usability Engineering: A List of Frequently Asked Questions*. 3rd edition. [online] Available from: http://www.ucc.ie/hfrg/resources/qfaq1.html [Accessed: 28/08/02]

Kirakowski, J. (2001) *Usability Evaluation by Questionnaire*. London, UK: Taylor & Francis Group

Klare, G.R. (2000) Readable Computer Documentation. In: *ACM Journal of Computer Documentation*, Vol. 24, Issue 3, pp. 148-168

Koffka, K. (1935) *Principles of Gestalt Psychology*. London, UK: Routledge

Köhler, W. (1947) *Gestalt Psychology: An Introduction to New Concepts in Modern Psychology*. New York, USA: Liveright

Koller, W. (1979) *Einführung in die Übersetzungswissenschaft*. Heidelberg: Quelle & Meyer

Koller, W. (1995) The Concept of Equivalence and the Object of Translation Studies. In: *Target*, 7 (2), pp. 191-222

Komissarov, V.N. (1977) Zur Theorie der linguistischen Übersetzungsanalyse. In: O. Kade (ed.) *Vermittelte Kommunikation, Sprachmittlung, Translation*. Leipzig: VEB Verlag Enzyklopädie, pp. 44-51

Korning Zethsen, Karen (1999) The Dogmas of Technical Translation – Are They Still Valid? In: *Hermes*, 23:65-75

Kussmaul, P. (1995) *Training the Translator*. Amsterdam & Philadelphia: John Benjamins

Kussmaul, P. (1998) Types of Creative translating. In: Andrew Chesterman (ed.) *Translation in Context*: Selected Contributions from the EST Congress Granada 1998. Amsterdam/Philadelphia: Benjamins, pp. 117-126

Landauer, T.K. (1995) *The Trouble with Computers: Usefulness, Usability and Productivity*. Massachusetts, USA/London, UK: The MIT Press

Larson, M. (1987) Establishing project-specific criteria for acceptability of translations. In: Marilyn Gaddis Rose (ed.), *Translation excellence: assessment, achievement, maintenance*. American Translators Association Scholarly Monographs, pp. 69-76. Binghamton, NY: University Center, State University of New York.

Larson, M. (ed.) (1991) *Translation: theory and practice, tension and interdependence*. Series American Translators Association Scholarly Monograph Series; Vol. 5. Binghamton, N.Y.: State University of New York at Binghamton

Lauscher, S. (2000) Translation Quality Assessment: Where Can Theory and Practice Meet? In: *The Translator* 6(2), pp. 149-168

Lee-Jahnke, H. (1998) Training in Medical Translation with Emhpasis on German. In: F. Massardier-Kenney and H. Fishbach (eds.) *Translation and Medicine*, American Translators Association Series, Vol. X, pp. 1-12

Lefevere, A. (ed.) (1992). *Translation / History / Culture: A Sourcebook*. London: Routledge, pp.141-165

Levý, J. (1967) Translation as a Decision Process. In: *To Honour Roman Jakobsen: Essays on the Occasion of his 70th Birthday,* Vol. 2, The Hague: Mouton, pp. 1171-1182

Linke, A., M. Nussbaumer and P. Portmann (1994) *Studienbuch Linguistik*. Tübingen: Niemeyer

Major, J.H. (1985) What Should You Write: A User's Guide, Tutorial, Reference Manual, or Standard Operating Procedure?. In: D. F. Beer (ed) *Manuals and Instructions: Giving Directions that Work*, IEEE. pp. 121-124

Mancuso, J.C. (1990) *Mastering Technical Writing*. Menlo Park, USA: Addison-Wesley Publishing Company

Markel, M. (2003). *Technical Communication*. Sixth Edition. Boston & New York: Bedford/St. Martin's

Mehlenbacher, B. (1993) Software usability: choosing appropriate methods for evaluating online systems and documentation. In: *ACM Proceedings of the 11th annual international conference on Systems documentation*. New York, USA: ACM Press, pp. 209-222

Microsoft Press (1998) *The Microsoft Manual of Style for Technical Publications - Version 3.0*. [CD-ROM] Redmond, Washington: Microsoft Press

Miller, G.A. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. In: *Psychological Review*, 63, pp. 81-97

Molich, R. and J. Nielsen (1990) Heuristic evaluation of user interfaces. In: *Proceedings ACM CHI'90 Conference on Human Factors and Computing Systems*. Seattle, USA: ACM Press, pp. 249-256

Mossop, B. (1998) The Workplace Procedures of Professional Translators. In: Andrew Chesterman (ed.) *Translation in Context*: Selected Contributions from the EST Congress Granada 1998. Amsterdam/Philadelphia: Benjamins, pp. 39-48

National Writers' Union (1998) Breaking into Technical Writing. [online] Available from: http://www.fredsampson.com/twcar03.html [Accessed: 16/05/04]

Newell A. and H. Simon (1972) *Human Problem Solving*. Englewood Cliffs, New Jersey, USA: Prentice-Hall

Newmark, P. (1988) A Textbook of Translation. Hemel Hempstead, UK: Prentice Hall

Newmark, P. (1990) *About Translation*. Clevedon: Multilingual Matters

Newmark, P. (1993) *Paragraphs on Translation*. Clevedon: Multilingual Matters

Nida, E. (1964) Toward a Science of Translating: with special reference to principles and procedures involved in Bible translating. Leiden: E. J. Brill

Nielsen, J. (2001) *First Rule of Usability? Don't Listen to Users*. [online] Available from: http://www.useit.com/alertbox/20010805.html. [Accessed: 29/06/02]

Nord, C. (1991) Scopos, Loyalty and Translational Conventions. In: Target 3(1), pp. 91-109

Nord, C. (1997) Translating as a Purposeful Activity. Manchester: St. Jerome Publishing

Norman, K. (2003) *Quant QUIS: Information About Quantitative Analysis*. [online] Available from: http://www.lap.umd.edu/QUIS/QuantQUIS.htm [Accessed: 19/09/2003]

Novak, J.D. (1998) *Learning, Creating , and Using Knowledge: Concept Maps™ as Facilitative Tools in Schools and Corporations*. New Jersey, USA: Lawrence Erlbaum Associates

O'Neill, M. (1998) Who Makes a Better Medical translator: The Medically KNowledgeable Linguist or the Linguistically Knowledgeable Medical Professional? A Physician's Perspective. In: F. Massardier-Kenney and H. Fishbach (eds.) *Translation and Medicine*, American Translators Association Series, Vol. X, pp.

Piaget, J. and B. Inhelder (1969) *The Psychology of the Child*. London, UK: Routledge & Kegan Paul

Pinchuck, I. (1955) *Scientific and Technical Translation*. London: André Deutsch

Power, R., D. Scott and A. Hartley (2003) Multilingual Generation of Controlled Languages. In: *Proceedings of EAMT-CLAW 03: Controlled Language Translation*. Dublin, 15th-17th May, 2003. Geneva: EAMT, pp. 115-123

Preece, J. (1993) *A Guide to usability: human factors in computing*, Wokingham, UK and Massachusetts, USA: Addison-Wesley

Preece, J. (1994) *Human-Computer Interaction*. Wokingham, UK and Massachusetts, USA: Addison-Wesley

Prescott, J. and M. Crichton (1999) Usability Testing: A Quick, Cheap, and Effective Method. In: *Proceedings of the 27th annual ACM SIGUCCS conference on User services: Mile high expectations*. New York, USA: ACM Press. pp. 176-179

Price, J. (1984) *How to Write a Computer Manual: A Handbook of Software Documentation*. Menlo Park, USA: The Benjamin/Cummings Publishing Company

Pym, A. (1996) Material Text Transfer as a Key to the Purposes of Translation. In: A. Neubert, G. Shreve and K. Gommlich (eds.) *Basic Issues in Translation Studies. Proceedings of the Fifth International Conference Kent Forum on Translation Studies II*. Kent, Ohio: Institute of Applied Linguistics, pp. 337-346

Quesenbery, W. (2001) *What Does Usability Mean: Looking Beyond "Ease of Use"*. [online] Available from: http://www.cognetics.com/presentations/whitney/more-than-ease-of-use.html [Accessed: 05/07/02]

Ramey, J. (1989) Escher Effects in Online Text. In: E. Barrett(ed), *The Society of Text: Hypertext, Hypermedia, and the Social Construction of Information*. Massachusetts, USA and London, UK: The MIT Press, pp. 388-402

Raskin, J. (2000) *The Humane Interface*. New York: Addison-Wesley

Redish, J.C. (1988) Reading to learn to do. In: Technical Writing Teacher, Fall, 15 (3), 223-233

Reiss, K. and H.J. Vermeer (1984) *Grundlegung einer allgemeinen Translationstheorie*. Tubingen: Niemeyer

Reiss, K. and H.J. Vermeer (1991) *Grundlegung einer allgemeinen Translationstheorie*. Tübingen: Niemeyer

Reiss. K. (1971) Möglichkeiten und Grenzen der Übersetzungskritik. Kategorien und Kriterien für eine sachgerechte beurteilung von Übersetzungen. Munich:Hueber

Robinson, D. (2003) *Becoming a Translator*. Second Edition. London & New York: Routledge

SAP (1997) *General Standards and Guidelines for Writing at SAP*. Walldorf, Germany: SAP AG

Savory, T.H (1957) *The Art of Translation*. London: Jonathan Cape

Schatzman, E.L. and F. Praderie (1993) *The Stars* (Translated by A.R. King). Berlin: Springer-Verlag

Schleiermacher, F.D.E. (1813) *Über die verschiedenen Methoden des Übersetzens* [On the different methods of translating]. Madrid: Gredos

Schneiderman, B. (1998) *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Massachusetts, USA: Addison Wesley Lingman

Schriver, K.A. (1997) *Dynamics in Document Design: Creating Texts for Readers*. New York, USA: John Wiley & Sons

Schveitser, A. (1987) *Übersetzung und Linguistik*. Berlin: Akademie

Schwitter, R., A. Ljungberg and D. Hood (2003) ECOLE: A Look-ahead Editor for a Controlled language. In: *Proceedings of EAMT-CLAW 03: Controlled Language Translation*. Dublin, 15th-17th May, 2003. Geneva: EAMT, pp. 141-150

Sherif, M. (1937) An Experimental approach to the Study of Attitudes. In: *Sociometry*, 1, pp. 90-98

Simpson, M. (1990) How usability testing can aid the development of online documentation. In: *ACM SIGDOC Asterisk Journal of Computer Documentation, Proceedings of the 8th annual international conference on Systems documentation*. Vol. 14, Issue 4. pp. 41-47

Smart, K.L. and M.E. Whiting (1994) Reassessing the documentation paradigm: writing for print and online. In: *ACM SIGDOC Proceedings of the 12th annual international conference on Systems documentation.* New York, USA: ACM Press. pp. 6-9

Sperber, D. and D. Wilson (1986) Relevance: communication and cognition. Oxford: Blackwell

Starr, M. (2001) *RU31: Give them Printed Documentation, too!!!* [online] Available from: http://www.tc-forum.org/topicus/ru31give.htm. [Accessed: 22/05/2004]

Steiner, G. (1975) *After Babel: Aspects of Language and Translation*. Oxford and New York: Oxford University Press

Sullivan, M. and A. Chapanis (1983) "Human Factoring a Text Editor Manual". In: *Behaviour and Information Technology, 2, 2*, pp. 113-125

Sykes, J.B. (ed.) (1971) *Technical Translator's Manual*. London: Aslib

Teague, R., K. De Jesus and M. Nunes Ueno (2001) Concurrent Vs. Post-Task Usability Test Ratings. In: *CHI '01 Extended Abstracts on Human Factors in Computing Systems*. pp. 289-290

Tebeaux, E. (1996) *The Emergence of a Tradition: Technical Writing in the English Renaissance, 1475-1640*. New York: Baywood Publishing

Tolman, E. (1948) Cognitive Maps in Rats and Men. In: *Psychological Review*, 55, pp. 189-209

Toury, G. (1995) *Descriptive Translation Studies and Beyond*. Amsterdam & Philadelphia: Benjamins

Treisman, A. (1960) Contextual cues in selective listening. In: *Quarterly Journal of Experimental Psychology*, 12, pp. 242-248

Treisman, A. (1964) Strategies and models of selective attention. In: *Psychological Review*. 78, pp. 282-299

Van Laan, K. and C. Julian (2001) The Complete Idiot's Guide to Technical Writing. Indianapolis, USA: Alpha Books

VDI 4500-2 - Technical documentation - Internal technical product documentation. Düsseldorf: VDI-Gesellschaft Entwicklung Konstruktion Vertrieb

Vermeer, H.J. (1978) Ein Rahmen für eine allgemeine Translationstheorie. In: *Lebende Sprachen*, 23(1), pp. 147-162

Vermeer, H.J. (1982) Translation als Informationsangebot. In: *Lebende Sprachen*, 27(2), pp. 97-l01

Vermeer, H.J. (1987a) What Does it Mean to Translate? In: *Indian Journal of Applied Linguistics*. 13(2), pp. 25-33

Vermeer, H.J. (1987b) Literarische Übersetzung als Versuch interkultureller Kommunikation. In: Alois Wierlacher (ed.) *Perspektiven und Verfahren interkultureller Germanistik*. Iudicium: Munich, pp. 541-549

Vermeer, H.J. (1989) Skopos and Commission in Translational Action. In: A. Chesterman (ed.) *Readings in Translation*. Helsinki: Oy Finn Lectura Ab, pp. 173-187

Vermeer, H.J. (1996) *A Skopos theory of translation (some arguments for and against)*. Heidelberg: TEXTconTEXT

Vygotsky, L.S. (1962) *Thought and Language* (E. Hanfmann & G. Vakar, Ed. & Trans.). Cambridge, Massachusetts, USA: The MIT Press

Weiss, E.H. (1985) *How to Write a Usable User Manual*. Philadelphia, USA: ISI Press

Wertheimer, M. (1959) *Productive Thinking*. New York, USA: Harper & Row

White, F.D. (1996) *Communicating Technology: Dynamic Processes and Models for Writers*. New York, USA: HarperCollins College Publishers

Witkin, H.A., P.T. Oltman, E. Raskin and S.A. Karp (1971) *Group Embedded Figures Test Manual*. Palo Alto: Consulting Psychologists Press

Wixon, D. and C. Wilson (1997) The usability engineering framework for product design and evaluation. In: M. Helander, T.K. Landauer, P. Prabhu (eds) *Handbook of Human-Computer Interaction*. 2nd edition. The Netherlands: Elsevier Science

Zirinsky, M. (1987) Usability Testing of Computer Documentation. In: *Proceedings of the 5th annual international conference on Systems documentation*. New York, USA: ACM Press, pp. 61-65

Appendices

Appendix 1: User Profile Questionnaire

User Profile Questionnaire

Name of Product:	
Name of Company / Organisation:	
Type of Company / Organisation:	
Number of Employees (where this product is used):	
Employee Job Title:	

Educational Background

1. What level of education is expected of employees?

☐ A. Degree ☐ B. Diploma ☐ C. Certificate ☐ D. Second Level ☐ E. Other

If you answered D or E, skip question 2.

2. What subjects were studied to obtain this qualification?

3. Which subjects previously studied were considered when recruiting employees?

Computer Skills

4. What level of computer skills are required for this job?

☐ A. Advanced ☐ B. Intermediate ☐ C. Beginners ☐ D. No Specific

5. Do employees typically have any specific computer qualifications prior to employent?

☐ A. Yes ☐ B. No

If yes, please give examples:

6. Do employees need these skills upon commencing employment?

☐ A. Yes ☐ B. No

7. What type of computer skills are needed for this job? Please tick all that apply.

☐ A. Word Processing ☐ B. Networking ☐ C. Email ☐ D. Databases

8. Any other computer skills?

Typing Speed

9. Is there a typing speed requirement for new employees?

☐ A. Yes ☐ B. No

10. If yes, what typing speed (in words per minute) is required?

☐ A. <20 wpm ☐ B. 20-40 wpm ☐ C. 40+ wpm

11. Please number the following skills in order of importance for potential employees?

☐ A. Typing Skills ☐ B. Computer Skills ☐ C. Language / Writing Skills

12. Are there any other skills which are vital for this job?

Appendix 2: Consent Form

Usability Study Consent Form

1. The Purpose of this Project

The purpose of this study is to examine a piece of software and its user guide. By participating in this study as a tester you will help us find ways of making software and user guides easier to use.

2. Your Role as a Tester

As a tester you will test the user guide to find problems and identify areas where we can make improvements. It is important to remember that you are not being tested – instead you will be testing the product and its user guide. If something is not clear, difficult or does not work, it means that there is a problem with the product and not with your abilities.

3. What will Happen?

The study will take place in a special usability laboratory in XXXXX where you will be observed using the product and user guide. You will be asked to spend 20 minutes familiarising yourself with the user guide before you perform a number of tasks using the software. When you have completed these tasks we will ask you a number of questions to find out your thoughts, opinions and feelings with regard to the user guide. One week after the test we will ask you to complete a short test to determine how well the user guide explained the information.

4. What Information will be Collected?

We will record information about how you use the product and the user guide. We will ask you to perform tasks and answer a series of questions in the form of a questionnaire. We will also record what happens on the computer screen as you perform tasks. In addition, all or some of the test will be videotaped. The purpose of this is to allow us to analyse the test in more detail.

5. Your Anonymity and Privacy

During the study you will be assigned a unique ID number. All information you provide, including documents, questionnaires, files and video tapes will be identified with this ID number only. With the exception of the test administrator, nobody will know your identity or be able to link your name to any of the data. The information you provide will be used in a PhD thesis and may also be used in research papers to be published in refereed academic journals. However, your anonymity will be assured and your identity will never be revealed.

6. Confidentiality

During the course of this study you will be using proprietary, copyright software which is supplied by a private company. Due to the commercial sensitivity of information relating to this software, you must not discuss the product or disclose any details relating to it. All documents and manuals must be returned to the test administrator before you leave the usability lab. You must delete all computer files relating to the product and return and print-outs made from these files.

7. Your Comfort During the Study

We will provide refreshments during the test and you may take a break at any time. Simply inform the test administrator that you would like to take a break. You are free to withdraw from the study at any time. If you have any questions regarding the study, please do not hesitate to ask the test administrator.

If you agree with these terms, please indicate your acceptance by signing below.

Name:	Test Administrator:
Signature:	Signature:
Date:	Date:

Appendix 3: Test Subjects Identity Form

Usability Study	Tester Details & IDs
Page 1 of 2	Created:
	Updated:

1. [name]

Tester ID Code	P1
Test Group	Control/Experimental
Email	
Signed Consent Form?	
Pre-Test Session (Date/Time)	
Test Session (Date/Time)	
Post-Test Session (Date/Time)	

2. [name]

Tester ID Code	P2
Test Group	Control/Experimental
Email	
Signed Consent Form?	
Pre-Test Session (Date/Time)	
Test Session (Date/Time)	
Post-Test Session (Date/Time)	

3. [name]

Tester ID Code	P3
Test Group	Control/Experimental
Email	
Signed Consent Form?	
Pre-Test Session (Date/Time)	
Test Session (Date/Time)	
Post-Test Session (Date/Time)	

4. [name]

Tester ID Code	P4
Test Group	Control/Experimental
Email	
Signed Consent Form?	
Pre-Test Session (Date/Time)	
Test Session (Date/Time)	
Post-Test Session (Date/Time)	

5. [name]

Tester ID Code	P5
Test Group	Control/Experimental
Email	
Signed Consent Form?	
Pre-Test Session (Date/Time)	
Test Session (Date/Time)	
Post-Test Session (Date/Time)	

6. [name]

Tester ID Code	P6
Test Group	Control/Experimental
Email	
Signed Consent Form?	
Pre-Test Session (Date/Time)	
Test Session (Date/Time)	
Post-Test Session (Date/Time)	

Appendix 4: Task Event Log

Usability Study	Task Event Log
Page 1 of 1	

Tester ID	
Date	

Time Measurements

Criterion	Instances		Total
1. Completion of Tasks	Task 1: Task 3:	Task 2: Task 4:	

Criteria Scores

Criterion	Instances	Total
1. Tasks Not Completed	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30	
2. Times user guide used	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30	
3. PTS Score	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30	
4. Incorrect icon choices	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30	
5. Incorrect menu choices	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30	
6. Verbal Interactions/Questions During Tasks	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30	
7. Observations of Frustration	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30	
8. Observations of Confusion	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30	
9. Observations of Satisfaction	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30	
10. Incorrect Commands / Input	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30	
11. Stopped Work without Completing Task	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30	

Individual Task Event Log

Appendix 5: Post-Test Survey

Usability Study	Post-Test Survey
Page 1 of 2	Tester ID: Date:

1. Which of the following statements is true?

a. QuicKeys are enabled if there is a checkmark beside the relevant menu item

b. QuicKeys are enabled if the relevant menu item is highlighted

c. QuicKeys are enabled automatically when you type a semi-colon

2. What happens if you type a colon after a QuicKey shortcut?

a. DigiLog will crash

b. The full name appears

c. Nothing happens

3. What happens if you use a QuicKey shortcut that is not recognised?

a. An error message appears

b. *XXXXX* appears in the log text

c. QuicKey creates a new shortcut for you

4. How do you change the appearance of the text in a log?

a. Go to the *Edit* menu

b. Go to the *Print* menu

c. None of the above

5. How do you align text in a log along the left side of the page?

a. Open the *Paragraph* menu and then click on *Format*

b. Open the *Format* menu and click on *Paragraph*

c. Open the *Setup* menu and click on *Paragraph*

6. The Setup menu is used to access which of the following?

a. QuicKeys

b. Text Formatting

c. Save Format

d. Display/Hide Toolbar

7. How do you display the status bar?

a. Click on the *Show/Hide Status Bar* icon

b. Click on *Status Bar* in the Setup menu

c. Click on *Status Bar* in the View menu

Usability Study | Post-Test Survey

8. How do you open an existing log?

a. Click on *Find* in the *File* menu

b. Click on *Open* in the *Setup* menu

c. Click on *Open* in *File* menu

9. How do you left align text in a log?

a. From the *View* menu, click on *Paragraph Format*

b. From the *Paragraph* menu, click on *Format* and then *Align Left*

c. From the *Format* menu, click on *Paragraph* and then *Align Left*

10. How do you paste text into a log?

11. What would you use the following toolbar button for?

a. To set the file format

b. To move a file to a different subdirectory

c. To change the size of text in a log

12. Which of the following are not menus in DigiLog?

a. File

b. Edit

c. Tools

d. View

e. Find

13. Which of the following file types *cannot* be created by DigiLog?

a. .HTML

b. .DOC

c. .RTF

14. Which of the following statements is false?

a. When you start DigiLog, a new log is automatically created

b. A new log is created when you go to the *File* menu and click *New*.

c. You create a new log by going to the *Setup* menu and clicking *New*.

15. How do you start DigiLog using the mouse?

Appendix 6: Results of Time Measurements

Participant	Task 1	Task 2	Task 4
P1	340	175	73
P3	355	210	41
P5	345	280	70
P7	352	442	136
P9	350	254	76
P2	225	110	64
P4	199	109	62
P6	208	83	72
P8	232	120	70
P10	214	112	65

Individual Task Times (in seconds) for Participants in Main Study

Group Averages	Task 1	Task 2	Task 4	All Tasks
Control	348.4	272.2	79.2	233.3
Experimental	215.6	106.8	66.6	129.7

Average Group Times (in seconds) for Each Task in Main Study

Appendix 7: Error Rates and Criteria

	Criterion	Control Group					Experimental Group				
		P1	P3	P5	P7	P9	P2	P4	P6	P8	P10
1	Tasks Not Completed	1	0	1	1	0	0	0	0	0	0
2	Times user guide used	2	4	3	6	3	2	1	2	3	2
3	PTS Score*	9.5	8.0	8.5	9.0	11.5	4.5	4.5	6.5	3.0	4.0
4	Incorrect icon choices	3	2	2	2	4	2	1	0	1	0
5	Incorrect menu choices	5	6	6	10	7	2	1	2	1	3
6	Verbal Interactions /Questions During Tasks	2	1	3	2	1	0	0	0	0	0
7	Observations of Frustra-tion	1	0	1	1	0	0	0	0	0	0
8	Observations of Confu-sion	1	3	4	3	2	0	1	1	1	0
9	Observations of Satisfac-tion (subtract)	0	0	0	0	0	0	0	0	-1	0
10	Incorrect Commands / Input	6	1	4	3	3	3	0	1	2	1
11	Stopped Work without Completing Task	2	0	0	0	1	0	0	0	0	0
	Individual Error Scores	**33**	**25**	**33**	**37**	**33**	**14**	**9**	**13**	**10**	**10**

Error Criteria Scores for Individual Participants in Main Study

	Criterion	Control	Experimental
1	Tasks Not Completed	0.6	0
2	Times user guide used	3.6	2
3	PTS Score*	9.3	4.5
4	Incorrect icon choices	2.6	0.8
5	Incorrect menu choices	6.8	1.8
6	Verbal Interactions/Questions During Tasks	1.8	0
7	Observations of Frustration	0.6	0
8	Observations of Confusion	2.6	0.6
9	Observations of Satisfaction (subtract)	0	-0.2
10	Incorrect Commands / Input	3.4	1.4
11	Stopped Work without Completing Task	0.6	0
	Group Error Scores	31.9	10.9

Average Error Scores for Each Group in Main Study

* PTS Score = Number of commands, icons & menus NOT remembered in Post-Test Survey

Appendix 8: QUIS Satisfaction Data

	Control Group					Experimental Group				
Question	P1	P3	P5	P7	P9	P2	P4	P6	P8	P10
3.1	6	5	5	5	4	5	6	6	7	6
3.2	5	1	5	3	3	4	7	6	7	7
3.3	6	5	7	6	6	7	6	8	6	5
3.4	5	1	5	4	4	4	5	6	7	5
3.5	3	2	5	3	6	7	7	4	5	6
4.1	6	1	8	6	7	6	4	6	8	4
4.11	7	1	5	7	7	7	6	3	8	4
4.12	6	1	9	7	7	5	6	6	8	6
5.1	6	5	7	7	5	9	7	9	6	8
5.1.1	6	5	5	7	7	7	6	8	8	7
5.1.2	7	7	5	7	7	9	9	8	6	8
5.2	9	7	5	6	6	8	9	7	8	8
5.2.1	4	1	6	7	6	8	5	7	7	8
6.1	7	1	5	6	4	5	8	5	8	7
6.1.1	4	1	7	3	6	4	5	7	7	8
6.2	6	5	7	7	6	7	9	7	8	8
6.3	6	1	5	5	4	6	7	7	9	7
6.3.1	9	3	4	5	4	5	7	7	6	5
6.4	6	1	7	8	4	8	9	4	8	6
6.4.1	7	2	5	4	3	7	8	6	7	7
6.4.2	7	3	4	8	5	8	7	7	6	7
6.5	3	9	3	6	4	7	8	6	3	7
7.1	9	2	6	6	8	7	7	7	8	8
7.1.1	3	3	7	7	7	7	4	6	7	7
7.2	9	2	7	7	6	6	6	7	9	7
7.2.1	5	3	4	7	2	5	8	7	8	4
7.2.2	7	2	5	9	6	6	9	8	8	7
7.3	5	1	3	6	5	5	7	8	5	6
7.3.1	9	1	8	5	4	7	7	7	6	7
7.3.2	7	6	5	5	9	7	9	6	9	9
7.4	4	1	5	6	4	5	9	7	9	7
7.4.1	3	6	4	4	4	5	7	8	7	5
8.1	5	3	5	6	4	6	9	8	8	5
8.1.1	9	2	4	6	6	7	7	8	8	6
8.1.2	4	1	5	4	4	6	8	7	5	5
8.1.3	4	2	3	6	5	6	9	7	5	5
8.2	2	5	5	7	6	4	9	4	5	9
8.3	4	7	6	7	7	6	6	8	6	5
8.4	6	3	5	7	6	6	6	7	7	6
8.4.1	9	1	5	7	5	7	7	7	6	9
8.4.2	4	1	5	7	3	7	4	8	7	7
Average Rating	5.839	2.926	5.390	6.000	5.268	6.292	7.048	6.707	6.975	6.536

Group Average: 5.08292683

Group Average: 6.71219512